APPLICATION OF ARTIFICIAL INTELLIGENCE MODELS IN TRAFFIC FLOW PREDICTION AND TIME-OF-DAY BREAKPOINTS DETERMINATION

BY

SYED MASIUR RAHMAN

A Dissertation Presented to the

DEANSHIP OF GRADUATE STUDIES

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the

Requirements for the Degree of

DOCTOR OF PHILOSOPHY

In

CIVIL ENGINEERING

MAY 2010

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS
DHAHRAN 31261, SAUDI ARABIA

DEANSHIP OF GRADUATE STUDIES

This dissertation, written by **SYED MASIUR RAHMAN** under the direction of his dissertation advisor and approved by his dissertation committee, has been presented to and accepted by Dean of Graduate Studies, in partial fulfillment of the requirements for the degree of **DOCTOR OF PHILOSOPHY IN CIVIL ENGINEERING.**

Dissertation Committee

Dr. Nedal T. Ratrout (Advisor)

Dr. Shokri Z. Selim (Member)

Dr. Hamad I. AbdulWahhab (Member)

Dr. Shukri H. Al-Senan (Member)

M. Mohandes

Dr. Mohamed Mohandes (Member)

Dr. Husain Jubran Al-Gahtani
(Chairman of the Department of Civil Engineering)

Dr. Salam A. Zummo
(Dean of Graduate Studies)

Date 21/6/10

DEDICATED TO MY PARENTS

ACKNOWLEDGEMENTS

First of all, the author is keen to express his humble, sincere gratitude to ALLAH, the most compassionate and the most merciful who gives him the opportunity to work and achieve the doctoral degree.

The author is very thankful and grateful to Dr. Nedal T. Ratrout for his continuous support, guidance, encouragement and advice during the whole period of doctoral program. The contributions, suggestions and comments of Dr. Shokri Selim, Dr. Hamad I. AbdulWahhab**,** Dr. Shukri H. Al-Senan and Dr. Mohamed Mohandes make him thankful to all of them.

The author is grateful to the Department of Civil Engineering and KFUPM for providing all sorts of support in conducting this research. The author would like to thank Advanced Traffic Analysis Center (ATAC) of North Dakota State University for providing data required for this study.

The author is greatly indebted to his parents, brothers, sisters and wife who continuously make supplication to ALLAH for making his work easy and smooth.

The author would like to express his gratitude to all of his friends and colleagues who generously have been supporting him from the beginning of his study in KFUPM.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

NOMENCLATURE

AD	Absolute Deviation
AI	Artificial Intelligence
AIC	Akaike Information Criteria
AIM	Abductory Induction Mechanism
ANFIS	Adaptive Neuro-Fuzzy Inference System
ANN	Artificial Neural Netwrok
App	Approach
AR	Autoregressive
ARIMA	Autoregressive Integrated Moving Average
ARMA	Autoregressive Moving Average
ATAC	Advanced Traffic Analysis Center
ATIS	Advanced Travel Information System
ATM	Advanced Traffic Management System
BPNN	Back-propagation Neural Network
CCF	Cross Correlation Function
CEM	Competitive Expectation Maximization
CFP	Cycle Flow Profiles
CO	Carbon Monoxide

CPM	Complexity Penalty Multiplier
DTA	Dynamic Traffic Assignment
EN	Expert Network
ENFTS	Ellipsoidal Neural Fuzzy Time-Series
ESM	Exponential Smoothing Method
FCM	Fuzzy C-Means
FIS	Fuzzy Inference System
FIS	Fuzzy Inference System
FLS	Fuzzy Logic System
FNM	Fuzzy-Neural Model
FRBS	Fuzzy Rule Base System
FRBS	Fuzzy Rule Base System
FSE	Fitting Square Error
GA	Genetic Algorithm
GMDH	Group Method of Data Handling
GMM	Gaussian Mixture Model
GN	Gate Network
GRNN	General Regression Neural Network
HOV	High Occupancy Vehicle

Hr	Hour
ITS	Intelligent Transportation Systems
KF	Kalman Filter
KF	Kalman Filter
K-NN	K-Nearest Neighbor
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MF	Membership Function
MNN	Modular Neural Network
MOE	Measure of Effectiveness
MOT	Ministry of Transportation
MSE	Mean Squared Error
MTL	Multitask Learning
NLPCA	Nonlinear Principal Component Analysis
O-D	Origin-Destination
PCA	Principal Component Analysis
PI	Performance Index
PSE	Predicted Squared Error
PSO	Particle Swarm Optimization

RAN	Resource Allocating Network
RBF	Radial Basis Function
REC	Regression Error Characteristic
RMSE	Root Mean Squared Error
RRBFN	Recurrent Radial Basis Function Network
SBC	Schwarz Bayesian Criteria
SC	Subtractive Clustering
STL	Single Task Learning
SVD	Singular Value Decomposition
SVM	Support Vector Machine
SVR	Support Vector Regression
TDNN	Time Delay Neural Network
TOD	Time-of-Day
UGPTI	Upper Great Plains Transportation Institute
UTCS	Urban Transportation Control System

ABSTRACT

NAME: **SYED MASIUR RAHMAN**

TITLE OF STUDY: **APPLICATION OF ARTIFICIAL INTELLIGENCE MODELS IN TRAFFIC FLOW PREDICTION AND TIME-OF-DAY BREAKPOINTS DETERMINATION**

MAJOR FIELD: **TRANSPORTATION ENGINEERING**

DATE OF DEGREE: **MAY, 2010**

This research aims to solve the problem of intersection traffic flow prediction along with the determination of time-of-day (TOD) break points for a pre-timed or actuated intersection. This study investigated mainly three artificial intelligence based models namely group method of data handling (GMDH) model, adaptive neuro-fuzzy inference system (ANFIS) model and type-2 fuzzy logic (FL) model to predict intersection traffic flow considering spatial and temporal characteristics. It is observed that that fuzzy c-means (FCM) clustering algorithm based ANFIS models and type-2 FL models outperform other considered models for two approaches of the considered intersection. Based on the obtained performance measures, it can be concluded that all the considered models are valid and promising for predicting traffic flow. This study also proposes a novel methodology in which time variable and Z-score of the approach traffic counts are used as the prospective features for determining TOD breakpoints instead of relying on the judgmental approach. The obtained results solve the problem of frequent changes of TODs. At the end, this study proposes a hybrid AI model consists of AIM, ANFIS, type-2 and artificial neural network (ANN) for predicting freeway traffic flow for the local freeway condition of Saudi Arabia.

الخلاصة

يهدف هذا البحث لحل مشكلة التنبؤ بحركة المرور بالإضافة إلى تحديد نقاط تقسيم حركة المرور خلال اليوم إلى أجزاء متجانسة (TOD) للإشارات الضوئية مسبقة التوقيت (Pre-timed) والمتفاعلة مع الحركة المرورية (Actuated). يقوم هذا البحث بشكل أساس على دراسة ثلاثة أنواع من النماذج المبنية على الذكاء الاصطناعي (AI) وهي: Group Method Data Handling (GMDH), Adaptive Neuro-Fuzzy Inference System (ANFIS) و (Type-2 fuzzy logic) للتنبؤ بتدفق حركة السير في التقاطع مع الأخذ بعين الاعتبار الخصائص المكانية والزمنية. وبناء على أداء هذه النماذج اتضح أن النماذج المستنبطة من نماذج (ANFIS) و (Type-2 fuzzy logic) تفوقت على باقي النماذج التي خضعت للدراسة, كما تبين أن كل النماذج التي أُخذت بعين الاعتبار فعالة وواعدة للتنبؤ بتدفق حركة السير. كما تقدم هذه الدراسة طريقة حديثة لتحديد نقاط (TOD) باستخدام قيمة (Z-Score) للحركة المرورية إضافة إلى الوقت كمتغيرين أساسيين, وقد حلت هذه الطريقة المشاكل التي تواجه عملية تحديد هذه النقاط في السابق, كما قدمت هذه الدراسة طريقة مستنبطة من نماذج الذكاء الاصطناعي (AI) للتنبؤ بالحركة المرورية على الطرق السريعة في المملكة العربية السعودية.

CHAPTER 1 INTRODUCTION

This research aims to solve the problem of intersection traffic flow prediction along with the determination of time of day break points for a pre-timed or actuated intersection. The background, problem statements, and objectives address both issues one after another for the sake of clarity. For the first part of this study, local data are not used due to the lack of readily available data and the methodology of this research is not constrained by any specific kind of traffic flow data. Local freeway data are used in the last part of this research in order to develop hybrid traffic flow prediction model.

Traffic flow is an important macroscopic traffic characteristic which plays an important role in planning, design, and operations of transportation facilities. The flow data are used for many traffic studies which include highway geometry, benefit of roadway improvements, estimates of road revenue, selection of highway routes, selecting the timing of maintenance, signal timing, location and design of highway systems, air quality analysis, design of traffic control systems and accident rates, average daily traffic, location of service areas (Federal Highway Administration, 2001). The traffic flow data can be measured through intrusive and non-intrusive methods. The intrusive methods consist of a data recorder and a sensor placing on or in the road, and non-intrusive methods are based on remote observation. The widely used intrusive methods are based on pneumatic road tubes, piezoelectric sensors and magnetic loops, and on the other hand, the non-intrusive methods mostly use manual counts, passive and active infra-red, passive magnetic, microwave radar, ultrasonic and passive acoustics, and video image detection for traffic counting. There are also advanced portable multiparametric detectors which count vehicles, observe speeds, obtain occupancy/density data, and determine vehicle types (Roess et al., 2004).

The traffic flow data collection in the location of an at-grade intersection is quite costly because only a count of typical four-leg intersection traffic flow with each movement classified by cars, trucks, and buses, requires the observation of 12 X 3= 36 separate pieces of data. The total costs for a roadside detector include capital costs (purchase and installation) and operational costs (maintenance, support and day-to-day operation). The capital cost of a detector ranges from $700 to $29,000, operation and maintenance cost varies from $100 to $2300, and the lifetime of the detector varies from 5 to 20 years (US DoT, 2007). In order to reduce cost of data collection, traffic flow studies of road networks can be conducted by adopting sampling procedures which relies on the assumption that entire networks, or specific sub portions of networks, have similar kind of demand patterns in time (Roess et al., 2004). This procedure can be implemented by conducting control and coverage counts. An at-grade intersection is one of the most complex location in a traffic system and it is difficult to determine representative control-count locations. Even it is not unlikely that the concerned network doesn't have any representative control-count location which will make the sampling procedure inaccurate and meaningless. In order to solve these kinds of problems intersection traffic flow prediction model can play a significant role. In this study artificial intelligence (AI) based models are proposed to predict intersection traffic flow in order to determine time-of-day (TOD) breakpoints although the predicted data can be used for many other purposes.

The recent research focuses on developing real-time adaptive signal control systems but the wide-scale implementation of such systems may be years away particularly in developing countries, due to the associated high costs for implementation and maintenance (Yin, 2008). Modern signal control systems are highly complex and are dependent on the sensor data supplied by intelligent transportation systems (ITS) but the basic forms of control such as TOD do not rely on the sensor data for operation (Hauser et al.,2000). TOD is defined as a mode in which two or more timing patterns are preset to match the traffic demands typically experienced at those times (Wang et al, 2005). In TOD mode, a day is segmented into a number of intervals such as the AM peak period, the PM peak period, nighttime in which the traffic volume patterns or demands are relatively stable in each segment (Wang et al, 2005). The lack of reliance on detectors is one of primary reasons for the popularity of TOD mode (Wang et al, 2005). The increased number of intervals will decrease the variation of the demand within each interval and the corresponding optimum traffic signal plan will better serve the traffic but the frequent changes in signal timings may be disruptive to the road users (Hua and Faghri, 1993).

In USA, the most widely used method for timing plan selection and implementation is TOD, where a pre-set plan is automatically used for a particular time interval (Gordon et al, 1996). In USA, basic forms of traffic control including pre-timed and actuated controls are widely used methods due to limited funding for the Department of Transportation and the difficulty in maintaining the sensors for support of advance control (Hauser et al., 2000). These circumstances pose the need of improved performances of widely used pre-timed signal control systems for both developed and developing countries. In pre-timed traffic signal control, the signal timings and cycle lengths may vary depending on TOD to reflect changes in traffic volumes and patterns. The actuated traffic control systems also require different traffic signal settings in order to cope with the changes of demand during different TOD because a single traffic signal setting doesn't perform well for an entire 24 hour period. The TOD mode can provide fairly efficient operation for both pre-timed and actuated traffic signal controllers provided that signal timing settings reflect current conditions. The proposed research is motivated to exploit the modern AI based techniques in predicting the traffic flow and subsequently determine TOD breakpoints.

Although there exist a number of optimization tools to assist traffic engineers in developing timing plans for a particular set of operating conditions, a few tools exist to help the engineer determine appropriate intervals, or to monitor an existing TOD system to ascertain if the conditions have changed sufficiently to require a new set of plans and/or intervals (Hauser et al, 2000). Usually, traffic engineers determine TOD breakpoints by analyzing 1 or 2 days worth of traffic data and relying on their engineering judgment. In fact, it is very difficult to determine optimum breakpoints unless traffic patterns change at certain times of the day and remain constant until the next change (Abbas, 2006). In fact, the proper selection of TOD breakpoints could improve day-to-day traffic operations which are only demonstrated in the recent literatures (Smith et al, 2001; Wang et al, 2005). But the number of TOD breakpoints and the duration of each interval should consider the effects of transition which mainly involves adjusting cycle length. However, very little research was found in the area of determining TOD breakpoints (Park et al, 2003a).

This study is proposing a new procedural model to determine optimum number of TOD breakpoints and the duration to avoid the reliance on judgment in determining TOD breakpoints. It will also provide the traffic volume which should be used for particular duration of TOD breakpoints. The prediction module of the proposed model will predict the traffic flow of all the approaches of the concerned intersections. The model will be developed based on the available data of a linear network in a city of USA.

The proposed framework will investigate the use of approach volume in determining the optimum TOD breakpoints. These points (i.e. TOD breakpoints) are the boundaries of the traffic conditions which are distinctive and require different traffic signal timing plans. The variations among approach volume are important parameters for determining traffic signal plans and consequently the TOD.

Therefore, the proposed study will enable the traffic engineer to use a systematic methodology to determine TOD breakpoints which will be optimum with respect to performance measures such as total delays, total stops, and total fuel used. It will eradicate the need of using existing judgmental approach in determining TOD break points. The existing engineering judgmental approach cannot identify the distinctive traffic conditions which require the implementation of different traffic signal timing plans to optimize the traffic flow throughout the 24-hr of a day.

1.1 Motivation

Although the recent trend of research focuses on developing real-time adaptive signal control systems, the wide-scale implementation of such systems seem to be years away in developing countries. The reasons mainly include associated high costs for implementation and maintenance, and dependency on the sensor data supplied by ITS but the pre-timed traffic signal control do not rely on the sensor data for operation. In pre-timed traffic signal control, the signal timings and cycle lengths may vary depending on TOD to reflect changes in traffic volumes and patterns. They can provide fairly efficient operation during peak periods provided that signal timing settings reflect current conditions. Unfortunately, there exists very little research in the area of determining TOD breakpoints to determine appropriate intervals, or to monitor an existing TOD system to ascertain if the conditions have changed sufficiently to require a new set of plans and/or intervals (see Hauser et al, 2000; Park et al., 2003a). These circumstances pose the need of improved performances of widely used pre-timed signal control systems. Even the actuated traffic control systems also require different traffic signal settings in order to cope with the changes of demand during different TOD because a single traffic signal setting doesn't perform well for an entire 24 hour period. But the determination of efficient TOD break points will heavily depend on reliable prediction of intersection traffic flow.

The proposed research is motivated to exploit the modern AI based techniques in predicting the traffic flow and subsequently determine TOD breakpoints which will provide a systematic procedural framework for the practitioners for both pre-timed and actuated traffic signal controls.

1.2 Objectives

The proposed research is intended to exploit the modern AI based techniques in predicting the traffic flow and determining TOD breakpoint of an intersection which is the middle one of a linear road network having three consecutive intersections. It will serve both pre-timed and actuated traffic control systems. The procedural technique can be achieved through the following objectives:

1. The first objective is to predict the approach traffic flow of an intersection of a linear network based on the observed historical traffic counts. The monthly average 24-hr traffic counts for each 15-min interval will be used for this prediction.
2. The second objective is to investigate a self-organizing model for intersection traffic flow prediction so that the practitioner can use the model without much intervention in the model building process. In this study, the group method of data handling (GMDH) algorithm based abductive network is considered as a self-organizing model.
3. The third objective is to use k-means and fuzzy c-means clustering model to determine TOD breakpoints.
4. The fourth objective is to find out design traffic flow data for the period of each TOD.
5. The fifth objective is to develop a methodology to evaluate the performance of TOD breakpoints with respect to a number of Measure of Effectiveness (MOE)s such as total delays, total stops, total fuel used, and CO emissions in microscopic traffic simulation environment.
6. The last objective is to investigate the potential of developing freeway traffic flow prediction model for the local context of Saudi Arabia.

1.3 Study Area

The study area consists of a linear road network of the City of Fargo, USA. The Advanced Traffic Analysis Center (ATAC) of the Upper Great Plains Transportation Institute (UGPTI) at North Dakota State University provides monthly average 15-min traffic counts of many locations of the City of Fargo (Figure 1-1). ATAC is collecting the data shown in the Figure 1-1 as a part of their commitment to establish a long-term program to support travel demand and transportation planning models, focusing on small-to-medium size urban areas such as Fargo. ATAC works with other relevant organizations to enhance modeling systems, improve supporting data, and to facilitate the use of models to support impact assessment from land-use or transportation network changes.

Three intersections along the 45th street are considered for the proposed study. The traffic counts are averaged for each month. The available traffic count includes twelve datasets of 24-hour traffic count. The complete datasets of the traffic counts can be accessed through the URL: http://www.atacenter.org. This study received approval of the Director of ATAC to use the data for research purpose. In this study, the traffic flow of the intersection crossing 45th street and 15th avenue will be predicted based on the past datasets of all three intersections. Figure 1-2 shows the simplified schematic diagram of the road which is considered in this study. Although there are a few other avenues in the study area, those avenues are neglected in this study. It is assumed that the proposed prediction models will perform well even without the data of those avenues.

For the last part of this study, local freeway data of Saudi Arabia are used to develop hybrid model. The description of the study area and the data description is provided in Chapter 6.

Figure 1-1: The location of automatic traffic count station at the city of Fargo, USA.

(Source: http://maps.google.com/)

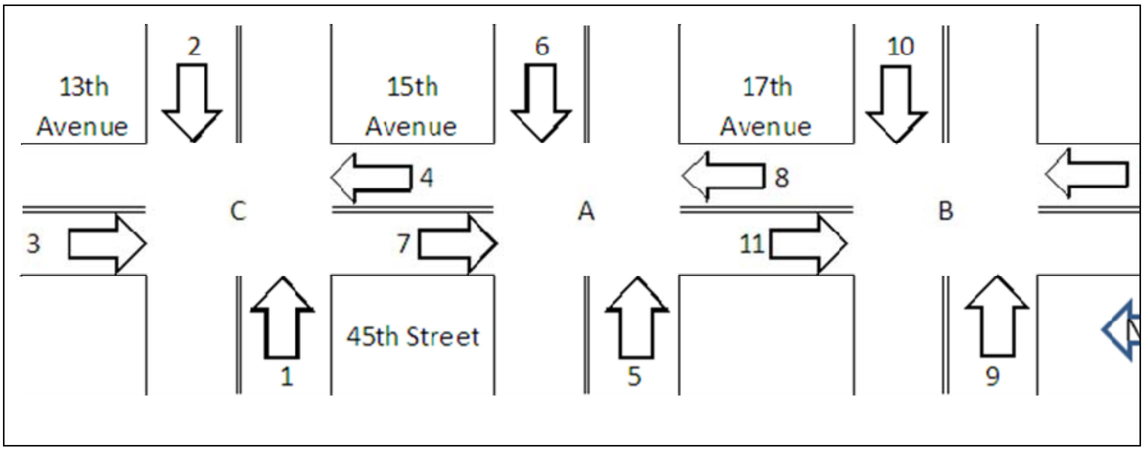


Figure 1-2: The number of approaches of the intersections along with 45th Street.

In order to perceive the considered intersections, their geometric features will be described. The geometric features are obtained by analyzing the detailed map available at the URL: http://www.bing.com/maps. At the intersection C, the 45th street has two exclusive left-turning lanes, three through lanes and one dedicated right-turning lane (Figure 1-3). The 13th avenue has two exclusive left-turning lanes, two through lanes and one dedicated right-turning lane (Figure 1-3). Actuated traffic signal controls are used in this intersection to control traffic flow. At the intersection A, the 45th street has one exclusive left-turning lane, three through lanes and one dedicated right-turning lane (Figure 1-3). The 15th avenue has one exclusive left-turning lane, one through lane and one dedicated right-turning lane (Figure 1-4). Actuated traffic signal controls are used in this intersection to control traffic flow. At the intersection B, the 45th street has two exclusive left-turning lanes, three through lanes and one dedicated right-turning lane (Figure 1-4). The 17th avenue has one exclusive left-turning lane, two through lanes and one dedicated right-turning lane (Figure 1-5). Actuated traffic signal controls are used in this intersection to control traffic flow.

Figure 1-3: The snapshot view of the intersection crossing 45th street and 13th Avenue.

(Source: http://www.bing.com/maps/)

Figure 1-4: The snapshot view of the intersection crossing 45th street and 15th Avenue.

(Source: http://www.bing.com/maps/)

Figure 1-5: The snapshot view of the intersection crossing 45th street and 17th Avenue.

(Source: http://www.bing.com/maps/)

1.4 Data Description

Any kind of numerical model-building requires understanding the inherent characteristics of the data. The description of the training and testing datasets is provided in Table 1-1. It indicates that 5 months data is used for training the model and 1 month data is used for testing the model. The data consists of 15-min traffic flow of all the approaches of the considered three intersections. A few statistical measures of the data are reported in Table 1-2. The means traffic flow varies between 10 and 150 vehicles/15 minute. The standard deviations of the data are high compared to the mean values which indicate higher variability of the data. Skewness is a measure of the asymmetry of the data around the sample mean and its negative value indicates that the data are spread out more to the left of the mean than to the right. On the other hand, the positive value of skewness means that the data are spread out more to the right. If the value of skewness is zero, it can be concluded that the distribution is normal distribution or any perfectly symmetric distribution. The skewness values of the data used in this study revealed that majority of the approach traffic flows are spread out more to the right of the mean and there is no clear indication that the data are generated from any perfectly symmetric distribution process. Kurtosis reveals the outlier-prone characteristics of a distribution and the kurtosis of the normal distribution is 3. The kurtosis values of the data indicate that all the traffic approach flows are less outlier-prone than the normal distribution.

Table 1-1: Summary of the available 15-min traffic counts of three intersections along with 45th street.

Number of Intersections	Interval of Traffic Count	Duration	Description of Data sets	Period	Number of Training Data Sets	Number of Testing Data Sets
3	15-min	24-hour	24-hour traffic flow for an average day of a given month expressed in 15-min intervals.	July, 2007-February, 2008	5X96 = 480	1X96 = 96

Table 1-2: Statistical measures of the traffic flow data of the concerned linear network.

Intersection	45 Street & 13 Avenue (C)				45 Street & 15 Avenue (A)				45 Street & 17 Avenue (B)			
Approach Number	1	2	3	4	5	6	7	8	9	10	11	12
Mean	142	159	174	128	13	17	149	146	75	58	149	139
Standard Deviation	100	120	116	93	12	15	101	105	55	45	105	110
Skewness	-0.05	0.12	-0.10	0.20	0.53	0.43	-0.13	0.18	0.15	0.31	0.11	0.57
Kurtosis	1.47	1.49	1.54	1.75	2.13	1.93	1.51	1.72	1.78	1.85	1.81	2.34

1.5 Organization of the Dissertation

This section provides the organization of the remaining chapters of this dissertation. Chapter 2 reports the literatures on traffic flow prediction models and TOD breakpoints. Chapter 2 also contains the literature on prediction models which are probably not yet implemented in the area of traffic flow modeling and prediction. Chapter 3 provides the methodology and describes the research to obtain traffic flow prediction of an intersection and to determine optimum TOD breakpoints. In chapter 4, three AI based models namely abductory induction mechanism (AIM) model, adaptive neuro-fuzzy inference system (ANFIS) model and type-2 fuzzy logic model are investigated to predict intersection traffic flow. Chapter 5 describes the clustering approach to determine TOD breakpoints and it also reports the performance of the TOD breakpoints obtained by using microscopic simulation software. Chapter 6 investigates the applicability of some advanced AI based models in predicting freeway traffic flow for the local condition of the Kingdom of Saudi Arabia. Chapter 7 provides the summary and conclusion of this dissertation and the prospective future research works relevant to this study.

1.6 Summary

This chapter describes the need of prediction of intersection traffic flow and the usefulness of the subsequent use of the predicted values for determining TOD breakpoints. This study proposes a novel approach for intersection traffic flow prediction by using GMDH based abductive networks which are not probably explored in this field. This model is self-organizing in nature which can be readily used by the traffic engineering professionals in the developing countries. It may be claimed that this study for the first time investigated the use of FCM based ANFIS models and singular value decomposition based type-2 fuzzy logic models for intersection traffic flow considering long time horizon. This research attempted to solve the prevailing judgmental approach in determining TOD break points. This chapter provides the statistical description of the used data which revealed that the data do not contain any outliers and they are generated from any symmetric distribution process. The majority of the Approach traffic flows are spread out more to the right of the mean. Finally, this chapter includes the organization of the dissertation.

CHAPTER 2 LITERATURE REVIEW

This chapter will encompass both the literatures on traffic flow prediction and the literatures on TOD breakpoints. Predictive traffic information can be categorized into strategic and short-term. This study requires traffic flow prediction of an intersection in order to determine TOD breakpoints which will enable efficient operation of pre-timed and actuated traffic controllers. But the current research on traffic flow prediction is dominated by short-term traffic flow prediction for adaptive traffic signal controllers. This research attempts to exploit the development in the field of short term traffic flow modeling and prediction research to predict intersection traffic flow which will be used to determine TOD break points. On the other hand, the research on the determination of TOD break points is not getting focus because its applicability is limited to pre-timed and actuated traffic signal controllers although most of the developing countries even the developed ones are relying on those types of traffic signal controllers heavily.

2.1 Traffic Flow Prediction Models

Traffic flow conditions can be divided into three traffic conditions which include free-flow, recurrent and non-recurrent congestion (Van Arem et al., 1997). When the demand exceeds the capacity of a road at a certain location then the traffic flow experiences recurrent congestion. On the other hand, incidents, maintenance works or special events cause non-recurrent congestion. Disbro and Frame (1989) determined the chaos behavior of traffic flow then many researchers including Roozemond and Johanns (1993) investigated the chaos phenomenon of traffic flow. They commented that certainly the chaos and fractals exist in traffic congestion state. The uncertainty exists in the traffic system which is a complex and huge system (Ming-bao and Xin-ping, 2008).

The traffic flow analysis can be divided into two main categories namely microscopic and macroscopic. Microscopic parameters are related to the behavior of individual vehicles in the traffic stream with respect to each other and macroscopic parameters describe the traffic stream as a whole (Quek et al., 2006). Microscopic analyses continuously or discretely predict the state of individual vehicles and primarily focus on individual vehicle speeds and locations. Macroscopic analysis aggregates the description of traffic flow, and the measure of effectiveness which are speed, flow and density (Boxill, and Yu, 2000). Microscopic parameters such as time and space headway of individual vehicles vary over a range of values depending on the speed of the traffic stream. In general, the microscopic parameters are related to the macroscopic flow parameters such as speed, volume, and density (Queck et al., 2006). Queck et al. (2006) suggested that microscopic analysis can be selected for moderate-size systems, where the number of vehicles is relatively small, and there is a need to study their individual behavior, and on the other hand, macroscopic analysis can be selected for higher density larger scale systems which only require a global behavior of groups of vehicles.

There are wide variety of traffic forecasting models which include nonparametric models (Smith, Williams and Oswald, 2002; Davis and Nihan, 1991), local regression models (Davis, 1990; Smith and Demetsky, 1997), neural network models (Zhong et al., 2005; Hu et al., 2008) neuro-genetic models (Abdulhai et al., 1999), neuro-fuzzy models (Park, 2002), genetic-neuro-fuzzy models (Hai-shuang et al., 2008), Markov chain models (Chen et al., 2005; Yu et al., 2003), Bayesian network models (Sun et al., 2006), spectral analysis based method (Nicholson and Swann, 1974). Hu et al. (2008) classified the research on short-term traffic flow forecasting into three categories: (1) Models based on statistic which include ARIMA model, time series models, Kalman filter theory and other nonparametric methods etc. (2) Traffic simulation based on dynamic network assignment or other theories. (3) AI based methods which include NN, SVR, Fuzzy logic based systems etc.

2.1.1 Need of Traffic Flow Prediction

The growing congestions on highways and surface streets became a worldwide concern for both developing and developed countries. In USA about three decades after World War II, it was aimed to accommodate growth by constructing new facilities which would have adequate capacity to handle future demand without regulating or controlling land use and economic growth (Giuliano and Wachs, 1992). During 1980's the policy goal of accommodating traffic demand started losing its feasibility due to the urban growth and automobile use, which has led to unprecedented levels of traffic congestion in many metropolitan areas (Hanks and Lomax, 1991). In order to address this congestion issue the transportation community can adopt three approaches which focus on increasing supply, managing demand and utilizing advanced technologies (Yang, 2006). In economics, demand is a function which refers to the relationship between price and consumption but transport planning often calculates demand at zero price, that is, free roads and parking (Litman, 2006). Transportation demand management and other similar approaches aim solving growth and traffic congestion problems by reducing or restricting travel demand (Ferguson, 2000).

The recent trend of transport development strategy is to meet the increasing demand with the help of advanced technologies such as ITS programs comprised of a number of technologies, including electronic surveillance, communications, information processing, traffic analysis and control technologies (Yang, 2006). ITS includes a myriad of products and services namely inter-modal transportation systems, intelligent traffic control systems, in-vehicle technologies, safety enhancement technologies, traveler advisory systems etc., and ITS reduces traffic congestion, improves road safety through intelligent traffic management (Shah and Dal, 2007). ITS benefits both transportation system users and mangers by improving the operational efficiency of transportation facilities, optimizing transportation management strategies and control actions, and disseminating current and future traffic information to both control centers and users to alleviate possible congestions through traffic redistribution (Yang, 2006).

Predictive traffic information can be categorized into strategic and short-term. Strategic information is typically needed for major decisions on road planning and includes prediction of traffic flows months or years into the future, and on the other hand, short-term predictive information often has a horizon of only a few minutes (Dia, 2001). Short-term forecasting of traffic conditions enables the modern transport management centers to proactively adopt appropriate control actions in advance in order to manage traffic more effectively (Yue and Yeh, 2008). The dissemination of this kind of information can also contribute in alleviating traffic congestions by redistributing traffic more evenly over road networks, and can enable travelers to reduce the costs of travel time and delays by selecting an alternative departure time or a route.

The short-term traffic forecasting problem "forecast the traffic condition in a link of interest (which will be referred hereby as the subject link) given past and /or current traffic conditions observed in the subject and upstream/downstream links" (Vythoulkas, 1993). The prediction of short-term traffic flow forecasting is an important component of advanced traffic management and information systems and it provides metropolitan traffic control centers with an automated tool for anticipating the congestion and its expected duration. This information can be provided to drivers in real time to give them realistic estimates of traffic conditions which are believed to have the potential to alleviate traffic congestion and enhance the performance of the road network (Dia, 2001).

ITSs are heavily dependent on traffic parameter predictions but the predictions should maintain the expected levels of accuracy regardless of the prevailing traffic flow conditions and especially in cases of extreme demand occurrence (Smith and Oswald 2003). Short-term traffic flow forecasting plays a very important role in ITS by supporting the development of traffic control strategies in Advanced Traffic Management Systems (ATMs), real-time route guidance in Advanced Traveler Information Systems (ATISs) and evaluation of dynamic traffic control and guidance strategies (Zheng et al., 2006). The dynamic traffic control and guidance systems require continuous traffic flow prediction for several minutes in the future using real-time traffic data (Cheslow et al. 1992). Without the ability to anticipate the short-term traffic conditions the full benefits of ITS cannot be realized (Smith et al., 2002).

In traffic modeling problems, the evolution of traffic flow in time and space considering complex settings such as freeway sections with variable geometric design, exits or entrances, or signalized streets with densely located signalization and uncontrolled mid-block demand, is still not clearly understood to develop accurate mathematical model (Vlahogianni, 2008). The peak traffic flow of a freeway network is very sensitive and slight losses of capacity can cause severe congestion with resultant costs in terms of user's time, fuel costs and pollution (Chen and Grant-Muller, 2001). A robust forecasting method capable of making accurate forecasts could have very real and substantial practical benefits if the method can be incorporated with freeway control system. The need of long term traffic flow prediction is already discussed in Chapter 1.

2.1.2 Causal Relationships among Traffic Flows

In this subsection, it is attempted to discuss the efforts in determining important issues related to the input selection process through spatio-temporal analysis, cross-correlation analysis, and principal component analysis.

Generally, traffic flow is a time-space phenomenon and the number and location of information sources effect the ability of a prediction algorithm to predict conditions accurately. According to Head (1995), the further away the used locations from the location of interest, the longer predicting horizons could be achieved but in urban areas the information from previous sites, located at far away is probably strongly distorted due to the effects of exogenous factors, such as traffic signals and traffic sources/sinks, and less reliable. Moreover, the temporal information (e.g. platoon dispersion) may become more distorted to contribute in building accurate model. There is a compromise between the distance among information sources and prediction accuracy (Head, 1995).

Vlahogianni et al. (2004) commented that in complex cases of traffic data with extreme fluctuation, the structure of forecasting algorithms must be adjusted tracing the temporal and spatial evolution of traffic characteristics which is usually accomplished by trying to forecast multivariate relationships or by using time and space-lagged input data. In this situation, the AI based models are preferred over the advanced form of many statistical methods and no-parametric regression. The AI based models can be multivariate in nature along with the ability of successfully absorbing the effect of missing or incorrect data and can adjust their internal adaptable form with the way traffic evolves.

In many studies data from upstream detector was considered for flow prediction. Park et al. (1998) investigated the correlation of traffic flow based on neighboring detector data and found that the correlation of traffic flow between upstream and downstream stations were strong. It suggests using neighboring detector data to improve traffic flow forecasting (Zhang and Ye, 2008).

In the literature, some models are proposed for traffic flow predictions which consider the spatio-temporal characteristics but rarely both spatial and temporal similarities of road links are considered (Hu et al., 2008). Sun et al., (2006) stated that some of the proposed models in the literature have not made use of information from the adjacent roads to analyze the trends of the object road and some approaches even didn't use the data at all. In order to forecast traffic flow precisely at least information from upstream and the subject road links are needed. Yin et al. (2002) proposed a fuzzy-neural model (FNM) to predict traffic flows in an urban street network based on only upstream flows. Jiang and Adeli (2005) proposed a nonparametric dynamic time-delay recurrent wavelet neural network model which incorporates the self-similar, singular, and fractal properties discovered in the traffic flow. The model used the autocorrelation function for selection of the optimum input dimension of traffic flow time series and incorporates both the time of the day and the day of the week of the prediction time. The model didn't consider the spatial characteristics of road links in building the model. Sun et al. (2006) proposed a Bayesian networks for traffic flow forecasting. In their model, traffic flows among adjacent road links in a transportation network are modeled as a Bayesian network and the joint probability distribution between the cause nodes and the effect node in a constructed Bayesian network is described as a Gaussian mixture model (GMM). The parameters of the GMM are estimated with the help of competitive expectation maximization (CEM) algorithm. The cause nodes are the data which are utilized for forecasting and the effect node are the data which are to be forecasted. But the model didn't consider the temporal characteristics of road links. Vythoulkas (1993) found out that the performance of a forecasting model is closely related to the location of upstream data collection stations, the use of the information of neighboring stations, and the time horizon of forecast.

Cross correlation function (CCF) is a standard statistics method to describe the statistical interdependence of two series which is used extensively in pattern recognition and signal detection, and seismology because it can precisely determine the differences in signal arrival times, or detect a known signal in a noisy one, or to search for cyclic date (Smith, 2003). The correlation coefficient is used to measure the relationship between two variables and the higher value of the coefficient generally indicates stronger relationship. If two data series are perfectly uncorrelated then the coefficient will become zero. Autocorrelation is a special form of CCF when bother series are identical.

The application of CCF should not used directly in determining cause and effect relationships. The correlation supports the notion of causation by providing clues regarding what the plausible causal relationships might be (Chen and Popovich, 2002).

For example, a perfect positive correlation cannot be directly considered as the causality of a relationship and on the other hand, a trivial correlation cannot be necessarily taken as an absence of causality (Yang, 2006).

Many researchers proposed a variety of multivariate approach of traffic forecasting by jointly considering the data from several locations. Stathopoulos and Karlaftis (2003) suggested investigating the varying influence of upstream data which may increase the accuracy of the model. Vlahogianni et al. (2003) used auto- and cross-correlated flow series in the NN model as external information and the obtained model performed considerably better than state-space and ARIMA models. Vlahogianni (2008) proposed genetically optimized modular neural network (MNN) architecture for predicting the anticipated traffic flow regimes. The model considers time-lagged volume and occupancy data, and information on the statistical characteristics of the short-term traffic flow evolution. The proposed MNN consists of an input layer in which the data are presented to the network and an output layer with four output neurons, one for each traffic flow regime.

Hu et al. (2008) used CCF to determine similarities between different traffic flow series. The most correlative road links and their time-delay are used constructing the Hybrid Process Neural Network which uses various scales to catch traffic features such as daily-periodicity, weekly-periodicity and spatio-temporal process. In this study, Wavelet transform was used to extract features of time series data on multi-scale and multi-resolution. He (2000) proposed Process Neural Network which relaxes the traditional neural network model of instantaneous synchronization input constraints. Time-varying process input signals, space aggregation operation, time aggregation operation, incentive threshold and incentive output composes a Process Neuron.

PCA is a modern tool in data analysis for extracting relevant information from confusing data sets (Shlens, 2009). It is a simple, non-parametric method which transforms a number of correlated variables into a smaller number of uncorrelated variables called principal components to reveal the sometimes hidden, simplified structures that often underlie within a complex dataset. The objective of PCA is to discover or to reduce the dimensionality of the data set and to identify new meaningful underlying variables. PCA provides the optimal information preserving transformation within the class of linear methods (Fukunaga and Koontz, 1970). It is reported in the literature that PCA facilitates many types of multivariate analysis by reducing dimensionality such as data validation and fault detection (Wise and Ricker, 1989), correlation and prediction (Joback, 1984).

Typical PCA is suitable for analyzing a two-dimensional matrix of data collected from a steady state process, containing linear relationships between the variables but these conditions are often not satisfied in practice (Bakshi, 1998). In order to address these limitations several extensions of PCA have been developed such as multiway PCA for analyzing multi-dimensional matrix (Nomikos and MacGregor, 1994), hierarchical PCA for easier modeling and interpretation of a large matrix by decomposing it into smaller matrices (MacGregor et al., 1994), dynamic PCA for extracting time-dependent relationships in the measurements by augmenting the data matrix by time lagged variables (Kresta et al., 1991), nonlinear PCA for extracting nonlinear relationships between the variables (Kramer, 1990), on-line adaptive PCA for updating the model parameters continuously by exponential smoothing (Wold, 1994).

Nonlinear PCA (NLPCA) involves nonlinear mappings between the original and reduced dimension spaces which ensure the more accurate description of the data if non-linear correlations exist between the variables (Kramer, 1990). The NLPCA method uses NN training procedures to generate nonlinear features by performing identity mapping, where the input is approximated at the output layer (Kramer, 1990). The proposed network of Kramer (1990) to determine NLPCA contains an internal layer containing fewer nodes than input or output layers, which force the network to develop a compact representation of the input data, and two additional hidden layers.

2.1.3 Parametric Time Series Model

In this subsection and the subsequent subsections, many available traffic flow prediction models are discussed. Over the last three decades, the researchers have been developing and applying new methodologies to forecast traffic flow in real-time by exploiting increased computational capabilities and availability of real-time ITS data (Castro-Neto et al., 2009).

Typically, the time series analysis focus on three basic time series attributes namely level, trend, and periodicity or seasonality. The relative magnitude of the time series values indicate the level which may vary around a constant mean or may increase or decrease with time. The changing level with time is referred as a trend which can be stepwise, linear, or non-linear. The presence of a pattern in a series which repeats itself at regular spaced intervals indicates the seasonality of the series. The time series models aim to account for and extract series trend and periodicity and then attempt to exploit the correlation structure of the de-trended and de-seasonalized data for best possible forecasting (Williams, 1999).

Box and Jenkins (1976) introduced the general model which includes autoregressive as well as moving average parameters, and explicitly includes differencing in the formulation of the model. The Autoregressive Integrated Moving Average (ARIMA) model is a generalization of an Autoregressive Moving Average (ARMA) model. The ARIMA model is one of the most mature theoretical and practical time series models and it is based on a systematic and iterative methodology but the fitting and maintenance of these models are time-consuming which impedes the real-time application (Zhi-Peng et al., 2008). The ARIMA models is typically defined as ARIMA(p,d,q) where, p is the autoregressive parameters, d is the number of differencing passes, and q is the moving average parameters.

Models can be formulated to determine theoretical or empirical relationships but any working physical model needs to be calibrated against data before making realistic forecasts (Van Arem et al., 1997). ARIMA type models including inputs from measurements at neighboring stations on the network, are the natural choice for linear physical relationships. Such models provide reasonable forecasts with a modicum of physical understanding such as "high flows upstream are likely to lead to high flows downstream after some delay depending on the distance after some delay depending on the distance between the stations" (Van Arem et al., 1997). It will work reasonably for free flow condition where the velocities can be assumed constant. In ARIMA type models the modeler can define the model form by specifying the range of stations to be included and the range of time lags.

ARIMA process is a commonly used model for short-term traffic flow prediction. The underlying assumption of ARMIA model is that traffic flow series can be made stationary by differencing and in most of the studies the investigated models were obtained by only first differencing (Yu and Zhang, 2004). ARIMA models are used for forecasting time series which can be made stationary by transformations which generally include differencing. Box and Jenkins (1976) provided an approach to systematically estimate ARMA models by identifying, estimating and diagnostic checking. But this approach relies on visual assessment of the correlation structure of training data. The problem of subjectivity is removed due to the combined effect of increased computing power, advanced software, and the use of information criteria such as Akaike information criteria (AIC), the Schwarz Bayesian criteria (SBC) for model selection (Han and Song, 2003). These information criteria provide a model selection metric which take care of the potential for overfitting (Smith et al., 2002).

Seasonal ARIMA model attempts to yield a stationary transformation for traffic flow series after normal and seasonal differencing (William, 1999). Der Voort et al. (1996) used unsupervised Kohonen self-organizing map to the cluster the traffic flow and then separate ARIMA(p, 0, q) model was fitted for each cluster. Yu and Zhang (2004) proposed a switching ARIMA model in which the patterns are the hidden states and they fitted a separate ARIMA model to each pattern of the traffic flow and apply the transition of hidden states to describe pattern changing. They found out that the patterns of traffic flow will last for some duration and the state is less likely to change for short duration and the state is more likely to change for longer duration. They introduced the variable of duration and used sigmoid function to model its influence to the transition probability.

The univariate time series models have been widely used for short-term traffic flow since the 1970s (Castro-Neto et al., 2009). Ahmed and Cook (1979) investigated Box-Jenkins model to develop a prediction model of the freeway traffic volume which was based on 166 data sets from three surveillance systems in Los Angeles, Minneapolis, and Detroit. The data sets were well-trained by an ARIMA model of order (0,1,3) but the moving-average parameters varied from location to location and over time. The models were more accurate for training datasets of freeway time-series data compared to the moving-average, double-exponential smoothing, and Trigg and Leach adaptive models with respect to MAE.

Levin and Tsao (1980) conducted Box-Jenkins time-series analyses were conducted on 20-, 40-, and 60-s volume data of two freeway locations and found out that ARIMA (0, 1, 1) model is the most statistically significant for all forecasting intervals for volume. The proposed model outperformed the Illinois Department of Transportation Traffic Systems Center ARIMA (0, 1, 0) model.

Hamed et al. (1995) attempted to develop time-series models for forecasting traffic volume in urban arterials based on Box-Jenkins approach using 1-min data set representing traffic volume on five major urban arterials. They found out that ARIMA model of order (0, 1, 1) is the most adequate model in reproducing all original time series.

Lee and Fambro (1999) investigated the use of the subset ARIMA model for short-term traffic volume forecasting. They identified four time-series models of different categories and used them for one-step-ahead forecasting task. The results reveal that the use of a subset ARIMA model could perform better in terms of accuracy of the short-term forecasting task compared to a full ARIMA model. Now-a-days, the univariate time series models such as ARIMA, exponential smoothing method (ESM), and double ESM are used for comparison purposes whenever a new forecasting model for short-term traffic is proposed (Park et al., 1998; Hansen et al., 2002).

ARIMA models are susceptible to missing or erroneous data which occurred due to malfunctions of sensors or data communication systems (Smith and Demetsky, 1995). Smith and Demetsky (1995) stated that traffic forecasts using time series were over predictive and lagging. This sort of characteristics of the models will make the prediction itself reactive in some sense (Abdulhai et al., 1999). Analyzing a few studies, Vlahogianni et al. (2004) concluded that the major deficiency of ARIMA models is their inclination to concentrate on the means and miss the extremes but in practice, traffic conditions show the opposite behavior with extreme peaks and rapid fluctuations. An efficient traffic forecasting model has to have the capability to capture the transition from stop-and-go situations to free flow conditions which is a shift from extreme values to smoother ones. Kirby et al. (1997) commented that the conventional time series approach takes no account of the structural relationship between a measurement at one point and another, in time which varies with traffic conditions; and the time can be estimated with the help of speed estimates. The previous measurements or the lags should not be regarded as fixed quantities rather that are related to the speed measurements. "Hence we see that, no other how well time series techniques might appear to fit the traffic data, their functional form is not necessarily one that is consistent with traffic characteristics" (Kirby et al., 1997).

According to the study of Hamed et al. (1995), the deficiencies of ARIMA models include the tendency of the model to concentrate on the means and miss the extremes, but traffic conditions exhibit the opposite behavior with extreme peaks. Now-a-days, the univariate time series model such as ARIMA is used for comparison purpose (Park et al., 1998; Hansen et al., 2002).

2.1.4 Exponential Smoothing Method

This method is widely used prediction technique to produce a smoothed time series and it requires little computation. The ESM assigns exponentially decreasing weights for the older measurements. The weights determine the degree of smoothing and how responsive the model is to fluctuation in the time-series data.

Exponential smoothing methods are simple, intuitive, and easily understood inexpensive technique which gives good forecast for wide variety of applications (Zhi-Peng et al., 2008). Due to the minimal requirements of data storage and computation this technique is suitable for real-time application.

According to the basic assumption of exponential smoothing the level of time series should fluctuate about a constant level or change slowly over time (Zhi-Peng et al., 2008) but this technique even in its adaptive form fail to give good forecasting when the time series takes on an obvious trend (Gui-Yan, 2004). The drawbacks of the Exponential Smoothing Method (ESM) are reflected in leading to large variations of states and determining an appropriate smooth constant. Zhi-Peng et al. (2008) proposed an improved adaptive exponential smoothing model (IAES) which consists of two models and one detector. One model is used to describe the pattern with level changing slowly and the other model is used to describe the pattern taking on strong trend, and the detector monitors the current pattern and determines which model is to be used for forecasting. Now-a-days, the univariate time series models such as ESM is used for comparison purposes whenever a new forecasting model for short-term traffic is proposed (Park et al., 1998; Hansen et al., 2002).

2.1.5 Non-parametric Regression

Non-parametric method is derived from nonlinear statistical theory which embraces a set of techniques for curve estimation without emphasizing on the true curve. In traffic forecasting, K-nearest neighbor (K-NN), kernel smoothing and local linear regression models are applied in which localized sets of traffic condition data are selected and utilized to generate the forecasts (Guo, 2005). In the current research K-NN is not used due to the required sizable historical database (Guo, 2005).

Non-parametric regression (NPR) technique does not depend on any rigid assumptions about the data and this technique searches a collection of historical observations for records similar to the current conditions which are used to estimate the future state of the system. It requires enough data to sufficiently describe the underlying process without using any prior knowledge about the system.

Davis and Nihan (1991) proposed a non-parametric regression method, the K-NN in order to overcome the inherent problems in parametric forecasting approaches. But the empirical study based on actual freeway data revealed that the K-NN method performed comparably to, but not better than, the linear time-series approach. They suggested focusing on the appropriateness of the forecast mean values for forecasting the extreme values characteristic of transitions from the uncongested traffic regime to the congested regime.

Smith and Demetsky (1997) developed historical average, time-series, neural network, and nonparametric regression models for the freeway traffic flow forecasting problem, which is defined as estimating traffic flow 15 min into the future. They concluded that the nonparametric regression model performs better than other models and experiences significantly lower errors than the other models. Moreover, the nonparametric regression model was easy to implement and proved to be have more generalization capability.

Smith, Williams and Oswald (2002) mentioned four difficulties in the implementation of NPR which include: choice of an appropriate state space, definition of a distance metric to measure nearness of historical observations to the current conditions, selection of a forecast generation method based on a collection of neighbors, and management of the potential neighbors' database. Shekhar (2004) proposed a NPR which is suitable for using in real-time system by reducing execution time with the help of advanced data structures and imprecise search. Shekhar (2004) also developed a methodology for implementing NPR.

2.1.6 Kalman Filter Model

Kalman (1960) proposed one of the advanced models in modern control theory which describes a recursive solution to the discrete data linear filtering problem. Kalman filter has been the subject of extensive research and application with the advances in digital computing. The ''state space model'' and the more widely known ''Kalman Filter model'' refer to the same basic underlying theory and typically, the term state space refers to the model and the term Kalman filter refers to the estimation of the state.

Generally, the process of estimating a state space models begins by estimating its ARMAX equivalent in order to capture the statistically significant AR and MA dimensions but the state space approach has advantages over the more widely used family of ARMA (ARIMA, ARMAX, etc.) models (Stathopoulos and Karlaftis, 2003). Stathopoulos and Karlaftis (2003) stated that the important advantages of Kalman filter models are both their explicit multivariate nature, which allows for data from different loop detectors to be jointly considered, and their ability to model a wide variety of univariate models, such as ARIMA, as special cases. They found out that the Kalman filter models are superior over a simple ARIMA formulation when modeling traffic data from different periods of the day. However, the Kalman Filer model is susceptible to produce high predictions or under-predictions when the traffic flow undergoes drastic changes (Zhang and Ye, 2008).

Okutani and Stephanedes (1984) proposed two models using Kalman filtering theory for predicting short-term traffic volume. The prediction parameters were improved using the most recent prediction error and the better accuracy in predicting the volume on a link is achieved by considering data from a number of links. Based on data collected from a street network in Nagoya City, the proposed models perform substantially (up to 80%) better than UTCS-2. Whittaker et al. (1997) proposed a dynamic state-space model for real-time modeling and prediction of motorway traffic. The Kalman filter provides the optimal state estimation for the proposed models.

In the recent literature, Kalman filter is getting used with other techniques to ensure better performance of the short term traffic flow prediction model. Stathopoulos et al. (2008) suggested a new AI-based approach in which a fuzzy rule-based system (FRBS) is augmented with an appropriate metaheuristic (direct search) technique to automate the tuning of the system parameters within an online adaptive rolling horizon framework and the proposed hybrid FRBS is used to nonlinearly combine traffic flow forecasts resulting from an online adaptive Kalman filter (KF) and an artificial neural network (ANN) model. Xie et al. (2007) investigated the application of Kalman filter with discrete wavelet analysis in order to address the problem of local noises. The decomposition of the original data by wavelet and subsequent use of Kalman filter model can help denoising data and increasing prediction accuracy. They found out that wavelet Kalman filter models based on Daubechies 4 and Haar mother wavelets outperform direct Kalman filter model in terms of mean absolute percentage error and root mean square error.

2.1.7 Artificial Neural Network Model

ANNs are biologically inspired systems consisting of massively connected processing elements analogous in functionality to biological neurons are organized in layers and are tied together with weighted connections corresponding to brain synapses. ANNs are generally designed by numerical-learning-based algorithms. The learning tools provide a given network with the capacity of adjusting its parameters in response to training signals. In supervised learning, by adjusting the weights of the network corresponding to a set of input and output exemplars, ANNs can be “trained” to approximate virtually any nonlinear function to a required degree of accuracy (Sadek, 2007). In unsupervised learning, the system is presented with a number of patterns and it is up to the network, through well-defined guidelines to categorize these patterns (Fakhreddine and De Silva, 2004). If sufficient numbers of hidden units (neurons) are available, then conventional ANN utilizing a single hidden layer and using arbitrary squashing functions can theoretically approximate any measurable function from a finite-dimensional space to another finite-dimensional space to any desired degree of accuracy (Hornik et al., 1989). Due to this capability, the ANN is regarded as a class of universal approximators.

The topology of ANNs refer to the ordering and organization of the nodes from the input layer to the output layer and the way the nodes and the interconnections are arranged within the layers of a given ANN determines its topology (Fakhreddine and De Silva, 2004). The selection of any topology depends on the type of concerned problems. Depending on the data processing nature, ANN topologies can be divided into feedforward and recurrent architecture. A network with feedforward architecture has its nodes hierarchically arranged in layers starting with the input layer and ending with the output layer, and in between, a number of hidden layers provide most of the network computational power (Fakhreddine and De Silva, 2004). The feedforward topology is associated with backpropagation learning (BPL) algorithm and is used by the multilayer perceptron network and the radial basis function network. On the other hand, the recurrent networks (RN) allow for feedback connections among their nodes and they are structured in such a way as to permit storage of information in their output nodes through dynamic states, hence providing the network with some sort of memory (Fakhreddine and De Silva, 2004). The Hopfield network and the time delayed neural networks (TDNN) have been designed based on the RN topology. The feedback connections in RNN enable them to handle problems involving dynamic processes and temporal patterns.

Due to the capabilities of self-learning and approaching any nonlinear functions, ANN is paid great attention to the system dynamic modeling (Xia and Wang, 2004). It has become widely used models for traffic flow prediction (Hu et al., 2008).The capability of real time implementation of ANNs for traffic flow forecasting is likely to be important for development and application of advanced traffic control in ITS. ANN prediction models have been found to provide significantly improved levels of accuracy when compared to the classical time-series prediction techniques (Smith, Williams and Oswald 2002). The research works on traffic flow prediction using ANNs are mainly focused on freeways and uninterrupted traffic flow (Dia, 2001; Ghosh, Basu and O'Mahony 2007; Xie, Zhang and Ye 2007).

Taylor (1994) proposed ANN model for freeway traffic data prediction. The weekday traffic flow was predicted by using the recent samples of traffic volume and occupancy at a station along with its upstream station data. Smith and Demetsky (1994) investigated the potentials of back-propagation ANN model compared to the traditional historical, data-based algorithm and time-series model. They stated that the ANN model was more responsive to dynamic conditions than the historical, data-based algorithm. It is aslo free from the lag and over-prediction characteristics of the time-series model. Ulbricht (1994) commented that recurrent ANN can solve the task of short-term traffic forecasting and they even outperformed the best results obtained with conventional statistical methods. Ulbricht (1994) investigated different architecture of ANN and finally obtained best results with proposed multi-recurrent network combining output, hidden and input layer memories with self-recurrent feedback loops of different strengths.

Mak (1995) proposed a Recurrent Radial Basis Function Network (RRBFN) for classifying and predicting temporal pattern. Mak (1995) tested two RRBFNs with different number of hidden nodes using a temporal sequence generated by an infinite impulse response filter and found out that the RRBFNs can approximate the filter more accurately than the continually running fully recurrent networks trained by the real time recurrent learning algorithm.

Park et al. (1998) used Radial Basis Function (RBF) NN for short-term freeway traffic volume forecasting. They compared the forecasting performance of RBF with the Taylor series, exponential smoothing method (ESM), double ESM, and back-propagation NN. RBF was performing better compared to others but the percentage error of the forecasted volume was 10 percent or more even for the best performance case.

Yasdi (1999) investigated the effectiveness of ANN system namely Recurrent Jordan Network for predicting time-series traffic volume data. They concluded that this type of NN has good generalization ability.

Chen and Grant-Muller (2001) proposed a dynamic NN based on a resource allocating network (RAN) which adapts a sequential learning scheme under which data points are presented to the network in sequence. The RAN is a single hidden layer network which starts with no hidden units. The allocation of a new hidden unit depends on the distance between the current input and the existing centers, and difference between the target and the output of the network. If both the distance and the difference are small then the existing parameters of the network can be adjusted by least mean squares (LMS) or extended Kalman Filter (EKF) algorithm to fit that pattern. A pruning scheme is also adopted to remove superfluous hidden units in the network.

Tan et al. (2005) put forward a traffic flow forecasting algorithm based on back-propagation (BP) NN for single road section which was followed with a grid computing model to meet the high-performance requirement of the forecasting process. Gird computing technology both integrates the grid resources and provides the traffic flow forecasting problem with resource sharing and coordination abilities, and helps solving the precious problems in traffic flow forecasting, such as low efficiency, low real-time, etc.

Pai et al. (2007) investigated the potential of ellipsoidal neural fuzzy time-series (ENFTS) model in predicting highway traffic which was originally developed for control and pattern recognition problems. The results of their model based on monthly traffic data at Tai-Shan tollgate of a freeway in Taiwan revealed that the ENFTS is superior to other neural network models, namely back-propagation neural networks (BPNN), and radial basis function neural networks (RBFNN) and general regression neural networks (GRNN) models.

Dongli et al. (2007) developed RBFNNs to predict the future behaviors of freeway traffic flow and proposed a Radial Basis Function neural network-based Model Predictive Control (RBF-MPC) for ramp metering. Through simulation results the authors concluded that the proposed approach can alleviate traffic jams and increase freeway capacity.

He et al. (2008) introduced process NN which has the capability to model spatio-temporal process into short-term traffic forecasting. The wavelet radix is used as weighted function expanding radix of process neurons to deal with the inputs on multi-scale. Through experimental results they claimed that the proposed model is suitable for real-time forecasting.

Tan et al. (2007) assumed traffic flow time series to be composed of a linear autocorrelation and a non linear structure. In the proposed hybrid approach, ARIMA model was used to predict the linear component of traffic flow time series and ANN were applied for the nonlinear residual component prediction. Lee et al. (2004b) proposed an ANN model that linearly combined the predictions from BPNN and RBFNN models using conditional probability and Bayes' rule.

Chen et al. (2005) proposed a new traffic volume prediction approach which was based on wavelet transform, ANN and Markov model. Initially, multi-resolution analysis was applied that is decomposition and reconstruction to the original traffic volume time-series to obtain a trend series and a hierarchy of detail series. The trend series was predicted through the training in ANN and Markov models were used to model and predict detail series. The proposed method was validated by a real traffic volume time series obtained in Suzhou city, China.

Jin and Sun (2008) proposed multitask learning (MTL) based NNs for traffic flow forecasting which addressed the limitations of single task learning (STL) models. The STL models do not take the advantage of the information provided by related tasks. Traditional forecasting methods are generally to anticipate with STL. The proposed MTL synchronal to train more than one task and can take full advantage of training information in the extra tasks to improve the generalization of the network which makes the net to have higher forecasting accuracy. In their experiments, they found out that the forecasted results of traffic flows are closer to the true value when MTL is used in the network.

The performance of ANN is influenced by the network training, the amount and quality of training data, network parameters such as the number of hidden layers, the transfer function, the number of epoch, the number of neurons in the hidden layers, the initial weights of the connection among neurons. Although ANN is effective for short-term prediction of traffic flow, it usually requires long training time (Zhao and Wang, 2007).

A conventional ANN works efficiently when the approximated function is relatively monotonic with only a few dimensions of the input features, but may not be as efficient for other cases (Haykin, 1994; Musavi et al., 1994). A monotonic function is either entirely non-increasing or non-decreasing and its first derivative (which need not be continuous) does not change sign. Generally, the ANN experiences the greatest difficulty in approximating functions when the input features are not linearly separable which implies that the approximated function has a higher complexity (Park D., et al., 1999). In a typical ANN, all input variables are connected to neurons of hidden layer which may affect the generalization ability. During training, if the value of some link weights are extremely greater than other weights then other links are omitted unintentionally (Afandizadeh and Kianfar, 2009). Generally, the researchers have to rely on time consuming and questionably efficient rules-of-thumb in developing optimum architecture of the NN (Vlahogianni et al., 2005).

2.1.8 Neuro-Fuzzy Model

Fuzzy logic theory allows the accurate representation of a given system behavior using a set of simple “if-then” rules but it is unable to tackle knowledge stored in the form of numerical data (Fakhreddine and de Silva, 2004). The rules for the mentioned kind of systems have to be extracted manually (Von Altrock, 1995). The knowledge about the systems in the form of linguistic and numerical data will make the problems even harder although this is the case for large-scale systems. ANN has been shown to be capable of learning virtually any smooth nonlinear mapping with a high degree of accuracy through a learning process in which numerical data are presented to the system for training under a computational structure composed of neurons and weighted links (Fakhreddine and de

Silva, 2004). The implicit representation of knowledge by ANN makes it difficult to explicitly explain the meaning of the weights among the nodes of the network once the systems have been trained. The decision making process is not explained very transparently by ANN (Goonatilake and Sukhdev, 1995).

In order to address the limitations of each system it is proposed to incorporate fuzzy logic reasoning within a learning architecture such as ANN. There are two major approaches in the mentioned hybrid system. One approach is concentrating on automating the generation of fuzzy rules using NN and optimizing the parameters of the fuzzy sets, and the other approach is tackling the issue of constructing ANN using fuzzy neurons (Fakhreddine and de Silva, 2004). A fuzzy neural network or neuro-fuzzy system is a learning machine that finds the parameters of a fuzzy system (i.e., fuzzy sets, fuzzy rules) by exploiting approximation techniques from ANN. Neuro-fuzzy systems which represent a class of hybrid intelligent systems combining the main features of ANN with those of fuzzy logic systems aim to circumvent difficulties encountered in applying fuzzy logic systems represented by numerical knowledge, or conversely in applying NN for systems represented by linguistic information (Fakhreddine and de Silva, 2004). Jang (1993) developed adaptive neuro-fuzzy inference system (ANFIS) which is described in Chapter 4.

Nauck and Kruse (1995) presented NEFCLASS, a neuro-fuzzy system for the classification of data which aimed to become interpretable in the form of linguistic rules and to be able to use prior rule based knowledge, so the learning has not to start from scratch. This approach is based on fuzzy perceptron which has the architecture of a usual multilayer perceptron, but the weights are modeled as fuzzy sets and the activation, output, and propagation functions are changed accordingly and it can be used to derive fuzzy neural networks or neural fuzzy systems for specific domains. Nauck and Klawonn (1996) showed that a combination of fuzzy clustering and neuro-fuzzy classification can produce better classification results, than either alone but the fuzzy sets obtained by projecting the clusters must be mapped to fuzzy sets used by NEFCLASS to initialized NEFCLASS.

Zhao and Wang (2007) developed a TSK-type fuzzy neural network model based on particle swarm optimization (PSO) to forecast the traffic flow of a city road intersection having traffic signal control. In this model, a fuzzy cluster module and several feed-forward neural network modules were comprised with a particle swarm optimization algorithm to optimize the structure and free parameters of the fuzzy neural network. The mean shift clustering algorithm generates the initial membership functions of input variables by determining their means and variances. Mean firing strength method is employed to remove the redundant fuzzy rules. Finally, PSO is used to fine tune the free parameters of fuzzy-rule neurons.

Yin et al. (2002) developed a fuzzy-neural model (FNM) which consists of a gate network (GN) and an expert network (EN) to predict the traffic flows in an urban street network. The GN finds out the clusters in the input data by adopting fuzzy logic based approach and the EN develops the input-output relationship like a conventional neural network. They proposed an online rolling training procedure which enhances the predictive power of the FNM through adaptive adjustments of the model coefficients in response to real-time traffic conditions.

Park (2002) proposed a hybrid neuro-fuzzy model for short term freeway traffic volume forecasting which consisted of a fuzzy C-means (FCM) method and a radial-basis-function (RBF) neural network. The FCM method classifies traffic flow patterns into a couple of clusters and RBF neural network develops forecasting models associated with each cluster. The study results showed that the proposed hybrid method is free from time-lag a phenomenon which is apparent in dynamic linear model and the RBF model without clustering.

Xiao et al. (2003) proposed a framework of a traffic prediction model that could eliminate noise caused by random travel conditions. The framework combined several AI technologies, such as wavelet transform, ANN and fuzzy logic, and the wavelet denoising method is emphasized.

Bao-ping and Zeng-qiang (2009) proposed a short-term traffic flow model based on ANFIS. The input data were obtained by a monitoring station at a section of freeway for a certain direction. They claimed that the training and checking errors are low enough to be accepted. Their proposed model showed good performance of simplicity, precision and stabilization. The input of the model consists of four immediate previous 5-min traffic flow data and the output is the next 5-min traffic flow data. They considered two Gbell membership functions which contain three changeable parameters.

Ming-bao and Xin-ping (2008) used subtractive clustering based Takagi-Sugeno (T-S) neuro-fuzzy model to develop the knowledge base of the traffic flow prediction model of chaotic time series. They used GA to find out the reasonable cluster radius for the subtractive clustering which provide the initial fuzzy rules and achieve the structure design of the knowledge base. GA was used offline in their study and its fitness function contains RMSE. They used Gaussian functions as the membership functions. They claimed that the model can be quickly trained online.

Quek et al. (2006) proposed a neuro-fuzzy model known as the pseudo outer-product fuzzy neural network using the truth-value-restriction method (POPFNN-TVR) for predicting short term traffic flow. In this POPFNN, the initial set of parameters and the system structure are constructed from a set of training data using an unsupervised learning algorithm and fine tuned with the help of the numerical information. The comparative results revealed that the prediction capability of the proposed POPFNN-TVR is higher than a conventional feedforward NN using backpropagation algorithm and it facilitated the understanding and analysis of traffic control problems by allowing a semantically rich set of rules to model system behavior.

An important issue in the design of neuro-fuzzy model is the identification of the fuzzy rules but there is no systematic design procedure at present (Quek et al., 2006). Quek et al. (2006) categorized the research works on generating and adapting fuzzy rules by using linguistic (Mamdani, 1975) as well as numerical information (Takagi and Sugeno, 1985). In one approach, linguistic information is used to identify fuzzy rules of the neuro-fuzzy model prior to the application of ANN techniques to adjust the rules (Keller et al., 1992; Yager, 1994). This subjective approach incorporate linguistic information rather than random choice of the initial fuzzy rules and it ensures faster convergence of the system during training and performs better. In the second approach, unsupervised learning algorithms are used to identify fuzzy rules of the neuro-fuzzy model before applying ANN techniques to adjust the rules (Lin and Lee, 1991; Ang and Quek, 2005; Hayashi and Nomura, 1992; Halgamuge and Glesner, 1994; Quek and Zhou, 1999). In this approach, the training dataset should be representative because it is the only source of information employed in this approach and the learning algorithms used for rule identification must be carefully selected in the absence of experts' opinions (Quek et al., 2006). In the third approach, supervised learning algorithm such as backpropagation technique is used to identify the fuzzy rules of the neuro-fuzzy model (Lee et al., 1994a; Ishibuchi et al., 1994; Yin et al., 2004). These models are essentially multilayered with the inputs and outputs as fuzzy membership values and the backpropagation learning algorithm is used to produce the mapping from inputs to outputs (Quek et al., 2006). These models appear as "black box" as the semantics of the structure remain opaque which contradicts the original intention for the integration of fuzzy concepts in neural networks.

2.1.9 Layered Models

Der Voort et al. (1996) termed the ATHENA and the KARIMA model as the layered models because these models "deal with non-stationary, cyclical nature of traffic flow series by first classifying the input series and then applying a class specific forecast model" (Williams, 1999). The French National Institute for Transport and Safety Research (INRETS) developed the ATHENA method which begins with a training set of historical data from the forecast location (Danech-Pajouh and Aron, 1991). The model first classify the current traffic volume curve based on Euclidean distance to the mean profiles which are calculated by taking average of the elements across all traffic flow curves in each class. A linear regression model is developed for each class to calculate the forecast volume. The Institute for Transport Studies at the University of Leeds in the UK developed the KARIMA model in which Kohonen self-organizing map is used for classification of traffic volume and different ARIMA models are selected for different clusters to forecast the traffic volume.

The combination of methods such as KARIMA and ATHENA destroy the inner relationship across time in the traffic condition series through clustering operation. In these approaches, the data at the boundary of each cluster may not be significantly different from each other while the forecasting strategies are totally different (Guo, 2005).

2.1.10 Support Vector Regression

Fundamentally, in Support Vector Regression (SVR) the data is mapped into a high-dimensional feature space via a non-linear mapping and then the linear regression is done in this space (Boser et al., 1992; Vapnik, 1995). Wu et al. (2008) estimated time lags between current and upstream traffic flow, and the most correlated upstream flow series used as the input of support vector regression model. The historical traffic flow of the concerned road was also used as input. The global SVR model was developed to make a whole day prediction. The Mean Absolute Percentage Error (MAPE) was used as an error measure to evaluate the performance of the global model. The duration during which the global model was not performing well, a local SVR model was proposed which was suitable for higher fluctuations of traffic flow. Ding et al. (2002) used four previous observations to predict the next observation of the traffic flow with the help of an ordinary SVR.

Qi-Hua and Rui (2005) proposed a traffic flow prediction model using support vector machines (SVMs). The model showed the capability of avoiding over-fitting and better generalization ability compared with BP or recurrent neural networks. Su et al. (2007)

proposed a short-term traffic flow prediction model and method based on online support vector regression (OSVR) according to the data collected sequentially by the probe vehicle or the loop detectors, which can update the forecasting function in real time via online learning way. The proposed model can be implemented as a real engineering application. The evaluation of the model by using I-880 database indicated that this model is superior to the back-propagation neural network (BPNN) model.

Zhang and Xie (2008) recommended v- support vector machine (*v*-SVM) model instead of ANN models for predictions of short-term traffic volume. Because v-SVM can address the problems of local minima and overfitting associated with neural network models. The test results revealed that for both one-step and two-step forecasting, the *v*-SVM model outperforms the multilayer feed-forward neural network (MLFNN) model for four data sets collected from three interstate freeways in terms of mean absolute percentage error and root-mean-square error.

2.1.11 Prediction Model for Real Time Traffic-Adaptive Signal Control

The real-time forecasting can significantly benefits the research on route guidance, incident management, public transportation, and traveler information (Perrin and Martin, 1998). The prediction of traffic flow plays a very important role in the development of real-time traffic-adaptive signal control. In early 1970s, the prediction of traffic flow was considered in the development of Urban UTCS system and it became a primary system component in the development of second generation (UTCS-2) and third generation (UTCS-3) control logic. The time horizon of predictions of demand for signal timing decisions was 5 to 15 min in UTCS-2 and approximately a cycle length in UTCS-3. Generally, increased time horizon for prediction will cause distortion of temporal information. Head(1995) mentioned three issues which are important for prediction traffic flow: (a) duration of the prediction time horizon, (b) number of prediction points per time horizon which is known as the prediction frequency, (c) number and location of information sources used for prediction.

The data-driven prediction techniques for urban signalized arterials have been used in many traffic control strategies, such as the UTC systems (Gartner, Stamatiadis and Tarnoff 1995) or Rhodes architecture (Mirchandani and Head 2001). Stephanedes et al. (1981) reviewed the UTCS and other demand predictors, and compared the prediction accuracy of UTCS-2, UTCS-3, historical averages, current measurement, and their proposed predictor having a parametric form similar to Proportional-Integral-Differential (PID) controller. The models were compared with respect to MSE and MAE for 5-min prediction and cycle-by-cycle predictions. They concluded that for 5-min predictions, the historical average performed better than UTCS-2 but both predictors were superior to the others.

Ledoux (1997) proposed cooperation based neural networks traffic flow model, which is intended to be integrated into a real time adaptive urban traffic control system. In this approach, the traffic flow is modeled on a single signalized link by a local neural network and the traffic flow is modeled over a wide network of junctions based on communications between local neural networks. This approach was based on simulated data generated in SSMT (Simulation Semi Macroscopique de Traffic) which was designed at Inrets in 1983 (Lebacque, 1983).

The deficiency in providing good temporally distributed prediction could be addressed by relying on actual flows which can be obtained by placing detectors on the links upstream from the intersection and use the flows at these points to provide predictions (Gartner, 1981). Real-time signal control systems such as SCOOT (Hunt et al., 1981), OPAC(Gartner et al, 1991), UTOPIA(Mauro and Taranto, 1990) adopted this approach in which the distance between the intersection and the upstream detector can constrain the prediction time horizon. $RHODES_{TM}$, another real-time traffic signal control system incorporates a traffic flow prediction algorithm namely PREDICT (Head, 1995). In this algorithm, the detector data on approaches of every upstream intersection, together with the traffic state and control plan for the upstream signals are used to predict future traffic demand. The algorithm is based on the assumption that all surrounding upstream intersections have fixed-time signalized planning which is violated in virtually every modern system (Zheng and Chu, 2008). In SCATS the detectors are located at the stop bars of the upstream intersection and the departure profiles along with a dispersion factor is used to predict the downstream arrivals. This algorithm considers the upstream signal in the prediction (Head, 1995).

2.1.12 Dynamic Traffic Assignment Based Model

Friesz and Bernstein (2000) stated that the traffic theory based methods emphasize on user equilibrium, or system optimal objective, or some variants of them. Those methods are categorized into Dynamic Traffic Assignment (DTA)-based approaches and non-DTA based approaches (Yang, 2006). The traffic assignment is based on the assumption of rational traveler behavior in route choice in terms of journey time, distance, monetary cost, congestion and queues, type of maneuvers required, type of road, scenery, signposting, road works, reliability of travel time and habit (Yang, 2006) but it is a difficult task to provide a generalized cost expression incorporating all these elements (Ortuzar et al., 2000).

Traffic dynamic data consists of the static demand data and the set of behavioral rules pertaining to the travelers' choices (Peeta and Ziliaskopoulos, 2001). DTA models are intended to perceive the interactions between traffic dynamics and driver behavior. Kaysi et al. (1993) stated that the behavioral models embedded in the DTA are more capable of realizing changes in traffic conditions over longer time horizons compared to the statistical methods which are only valid for relatively short forecasting methods.

In the analytical category, mathematical programming DTA models formulate the problem in a discretized time setting, and the optimal control methods formulate it in a continuous setting but both the methods suffer from many limitations to handle realistic traffic scenarios (Yang, 2006). Simulation-based models address the inability of analytical representation of traffic flow by replicating dynamic traffic phenomena and by capturing the complex vehicle interactions. The simulation based models can be classified into macroscopic, meso-scopic, and microscopic ones.

The research of Lighthill and Whitham (1955) and Richard (1956) introduced the dynamic and macroscopic modeling of traffic flow, a simple first-order continuum model which requires time-varying O-D matrices. The research works of Ashok and Ben-Akiva (1993, 2000), and Hellinga (1997) emphasized on the estimation of time-varying O-D matrix. Papageorgiou (1998) commented that it was unlikely for any macroscopic traffic flow theory to reach the descriptive accuracy for other domains.

2.1.13 Evolutionary Fuzzy Model

In this subsection and the subsequent subsections, the prediction models which are mainly successfully implemented in the areas other than the traffic flow prediction are going to be reviewed.

Thrift (1991) investigated the applications of a genetic algorithm (GA) for optimizing the behavior of fuzzy systems. Evolutionary algorithm solves the problem of local minima and provides number of solutions at a time. The tuning of fuzzy system is done by optimizing its parameters which include the shape of fuzzy set and its parameters, overlapping and parameter of aggregation operators (Chaturvedi, 2008). GA is a well known and widely used global search technique with the ability to explore a large search space for suitable solutions only requiring a performance measure and the generic code structure. The independent performance features of GAs make them suitable candidates to incorporate a priori knowledge for fuzzy rule based systems (FRBSs) in the form of linguistic variables, fuzzy membership function parameters, fuzzy rules, number of rules, etc (Herrera, 2008). Due to these capabilities extended GAs are used in the development of a wide range of approaches for designing FRBSs over the last few years.

Fuzzy systems have been proven successful for many applications of classification, modeling, and control problems by mainly incorporating human expert knowledge (Cordon et al., 2004). But the lack of learning capabilities characteristics of fuzzy systems encouraged the researchers to solve the problem by introducing hybridization. Herrera and Verdegay (1996) provided a hybrid model of Genetic Fuzzy Systems (GFS) to add learning capabilities. A GFS is the fuzzy system which is augmented by a learning process based on a GA (Cordon et al., 2004). There are three main approaches for learning rules of a rule based systems which are Pittsburgh approach (Smith, 1980), Michigan approach (Holland and Reitman, 1978), and iterative rule learning approach (Venturini, 1993). In Pittsburgh approach the entire rule set is represented as a chromosome and in Michigan approach chromosome represents individual rule and a rule set is represented by the entire population. In the iterative approach, chromosomes code individual rules, and a new rule is adapted and added to the rule set iteratively.

As an alternative of fuzzy neural networks and fuzzy decision tree induction, Yuan and Zhuang (1996) proposed a Fuzzy Genetic Algorithm (FGA) to generate fuzzy classification rules. They also used multi-value logic coding, composite fitness function, viability check, and rule extraction to improve the efficiency and the effectiveness of the algorithm.

GA can be used “from the simplest case of parameter optimization to the highest level of complexity of learning the rule set of a rule based system” (Cordon et al., 2004). GA performs parameter optimization which is typically used in genetic fuzzy clustering and it can also learn or tunes different components of fuzzy rule-based systems (GFRBs) of genetic fuzzy rule-based systems (GFRBSs). GA can circumvent the problems of spurious rules in a fuzzy model by determining membership functions with a specified number of fuzzy rules, finding fuzzy rules with known membership functions, and finding both membership functions and fuzzy rules simultaneously (Mitra and Hayashi, 2000).

Ju et al. (1998) proposed a Genetic-Based Fuzzy Model (GBFM) to process Korean financial data and modeling time-series process by concisely representing the fuzzy rules with one or more FAM (Fuzzy Associative Memory) matrices. Buckley and

Hayashi (1994) presented fuzzy genetic algorithms to (approximately) solve fuzzy optimization problems, the fuzzy maximum flow problem, fuzzy regression, and tuning a fuzzy controller problems. Damousis and Dokopoulos (2001) present a fuzzy expert system for forecasting the wind speed at a wind energy conversion system (WECS) site and the generated electrical power. The implementations of two genetic algorithms were used for the training of the expert systems. Muhammad and King (1997) proposed an evolutionary fuzzy network method for predicting financial exchange market which was used in the Santa Fe time series forecasting competition, 1990-91. They used GA to adapt the parameters of the fuzzy network in order to obtain the best performance.

Kim (1997) proposed a genetic fuzzy predictor ensemble (GFPE) model for time series prediction by adopting two design stages, where the first stage generates a fuzzy rule base that covers as many of training examples as possible and the second stage builds fine-tuned membership functions that make the prediction error as small as possible. These two design stages were repeated independently for different partition combinations of input–output variables and finally, the GFPE combines multiple fuzzy predictors by an equal prediction error weighting method. The proposed model was applied on both the Mackey–Glass chaotic time series and the non-stationary foreign currency exchange rate prediction problem.

Yang and Huang (1998) proposed a new self-organizing model of fuzzy autoregressive moving average with exogenous input variables (FARMAX) for one day ahead hourly load forecasting of power systems by identifying the fuzzy model as a combinatorial optimization problem. An evolutionary programming (EP) scheme is used along with heuristics to determine the optimal number of input variables, best partition of fuzzy spaces and associated fuzzy membership functions.

Hansen et al. (2002) proposed ANN whose architecture was determined by genetic algorithm in order to predict time series. Their model outperformed ARIMA models in six different time series examples and ANNs detect patterns in data which remain hidden to the autoregression and moving average models.

Recently, the integration of ANN and fuzzy set theory approach has been proposed for many different applications in order to provide improved results by exploiting the benefits of both the approaches (Chaturvedi, 2008). The typical neuron model is modified to obtain a GN model using fuzzy compensatory operators as aggregation operators to solve the problems of large number of neurons and layers required for complex function approximation, which affect both training time and tolerant capabilities of ANNs (Chaturvedi, 1997). A GN model can be developed by using fuzzy compensatory operators that are partly union and partly intersection which combine summation ($\sum$) and product (π) neurons at different layers. Typically, the number of weights of a GN model is very low compared to a multi-layer feedforward network. It contributes in reducing training time and the required number of training patterns. GN model is less complex as compared to multi-layered feedforward ANNs and is more flexible at structural level (Chaturvedi, 2008).

2.1.14 Evolutionary Neural Network

Evolutionary Computation is a collection of stochastic optimization algorithms which is loosely based on the concepts of biological evolutionary theory (Vonk et al., 1997). Genetic Algorithms are a kind of EC which are successfully used in many applications including the optimization of neural network architecture. Holland developed GAs in the 1970's which are based on a Darwinian-type survival of the fittest strategy in which stronger individuals in the population have a higher chance of creating offspring (Holland, 1975). Each individual in the population represents a potential solution and they are made up of a set of genes and in the traditional case of binary strings these genes are just bits. GAs intend to determine the individual with a maximum fitness value by using a stochastic global search technique which works in the solution space. The operation of a standard GA is as follows:

1. randomly generate an initial population of chromosomes
2. compute the fitness values of each individual of current population
3. make an intermediate population by extracting individuals out of the current population by using reproduction operator
4. generate the new population by applying genetic operators such as crossover, mutation to this intermediate population
5. if an individual of the current population satisfies the problem requirement then stop, otherwise go to step 2

   (Vonk et al., 1997)

Evolutionary computation can be used in the field of ANN to train, analyze and generate the architecture of an ANN and its weights (Vonk et al., 1997). GAs do not use error gradient information which means that they can be used where this information is not available and the activation function of the neurons doesn't have to be differentiable or even continuous. In principle, GAs can be used to train any type of neural network including recurrent networks (Vonk et al., 1997).

GA is used to optimize the weights of a NN by performing global search of the weight space and unlikely get stuck in a local minima. The alternative back propagation (BP) learning algorithm often gets stuck in local minima. It doesn't use error-gradient information and it does not require the activation function of the neuron to be differentiable or continuous. However, the use of BP to fine-tune a near optimal solution obtained by GA is reported to be successful (Hibbs, 1994). In general, GA can be used in the field of NNs to train the weights of a NN, to analyze a neural network, to generate the topology of a neural network and to generate both the topology and weights of the neural network (Vonk et al., 1997).

Hassoun (1995) used GA on a fixed three layered feedforward NN to find the optimal hidden targets from the input. Eberhart (1992) used GAs to find the input patterns that yield a certain output of the NN. Braun and Weisbrod (1991) used GA to generate feedforward NNs using a direct encoding scheme the genes of a chromosome represents a connection between two neurons. Jacob and Rehder (1993) represented a NN topology as a string consisting of all the existing paths from input to output neurons in the GA. Different GAs were used for creating topology, activation functions and weights as separate modules.

Vlahogianni et al. (2005) used GA based, multilayered structural optimization strategy in the proper representation of traffic flow data with temporal and spatial characteristics and in the selection of the appropriate NN architecture. They concluded that the capabilities of a simple static NN, with genetically optimized step size, momentum and number of hidden units, are very satisfactory for modeling both univariate and multivariate traffic data.

Zhong et al. (2005) used genetic algorithms to design time delay neural network (TDNN) models as well as locally weighted regression models to predict short-term traffic for six rural roads from Alberta, Canada which are from various trip-pattern groups and functional classes. Refined regression models showed higher accuracy compared to refined TDNN models for both stable and unstable traffic flow conditions.

GA can contribute in decreasing the model complexity and improve model generalization ability by determining optimum layout of hidden layer connection links (Afandizadeh and Kianfar, 2009). Afandizadeh and Kianfar (2009) proposed a hybrid model in which an initial model is synthesized by NN and then the model structure is optimized through GA to improve the generalization capability. The fitness function of the GA considered MSE of the network for validation data.

2.1.15 Genetic Neuro-Fuzzy Model

GA, fuzzy logic (FL), and ANN are frequently used artificial intelligence (AI) techniques which are complementary rather than competitive. Many researchers have hybridized GAs, FL, and NNs to develop a better performance model but most hybrid models use a multistage combination or identify partial parameters required in the model resulting in sub-optimal solutions (Cheng and Ko, 1996). The fusion of those models provides an alternative to a strictly knowledge-driven reasoning system or a purely data-driven one and learned that the combined models can provide a more accurate and robust solution than can be derived from any single technique because typically the methods do not try to solve the same problem in parallel but they do it in a mutually complementary fashion (Shapiro, 2002).

Ishigami et al. (2000) presented an auto-tuning method of fuzzy inference using GA but the determination of membership functions of the fuzzy inference typically depends on human experts, which is a difficult problem and time-consuming. They proposed some auto-tuning methods to reduce the time-consuming operations with the help of an auto-tuning method for the fuzzy neural network by GA. Farag et al. (1998) proposed a fuzzy-neural network (FNN) model which can handle both quantitative (numerical) and qualitative (linguistic) knowledge. In their model, a multi-resolutional dynamic genetic algorithm (MRD-GA) is used for optimized tuning of membership functions of the model.

Ichimura and Tazaki (1995) proposed reasoning and learning method for fuzzy rules using ANN with the adaptive structured GA that determines the neural network structures and their input weights. The adaptive structured GA can generate or annihilate the specified units respectively in hidden layer to achieve an overall good system without requiring a priori assumption about network topology but it requires only the input and output characteristics of the task. Cheng and Ko (1996) fused GA, FL, and ANN to develop an evolutionary fuzzy neural inference model (EFNIM) that uses GA to simultaneously search for all parameters required in fuzzy neural networks (FNNs) by encoding variables in FL and NNs with the help of two approaches, summit and width representation method (SWRM) and block-representation method (BRM). Balkatr (2002) proposed an evolution-based approach to design of neural fuzzy networks which optimizes the whole fuzzy system with minimum rule number according to given specifications, while training the network parameters. The approach combines evolution strategies and simulated annealing algorithms in finding the global optimum solution and the optimization variables include membership function parameters and rule numbers which are combined with genetic parameters to create diversity in the search space due to self-adaptation. The optimization technique doesn't depend on the considered topology and capable of handling any type of membership function.

Dash et al. (2000) proposed a self-organizing fuzzy-neural network with a new learning algorithm and rule optimization using GA. The adaptive model performs much better than the existing ANN techniques for predicting electrical load in a time frame varying from 24 to 168 hours based on the practical data collected from the Virginia Utility, USA. Huang et al. (2001) proposed an integrated neural-fuzzy-genetic-algorithm (INFUGA) to predict permeability from well logs in a petroleum reservoir in Australia by following five steps: (1) select appropriate well data sets; (2) generate fuzzy rules by neural networks; (3) generate hyper-surface membership functions by neural networks; (4) optimize defuzzification operator parameters by genetic algorithms; and (5) interpolate fuzzy rules to provide estimates. It is an assumption-free, model-free, and adaptive estimator and is suitable for handling multi-dimensional inputs and outputs without requiring a structured knowledge base but it requires additional CPU time for the performance.

Farag et al. (1998) exploits the inherent benefits of the fuzzy logic theory, neural networks, and genetic algorithms (GAs) to develop a fuzzy-NN model which can handle both quantitative and qualitative knowledge. The learning algorithm of the model includes finding the initial membership functions of the fuzzy model, extracting the linguistic-fuzzy rules, and optimizing membership functions of the model with the help of a multi-resolution dynamic genetic algorithm (MRD-GA).

Abraham (2002) commented that in an integrated neuro-fuzzy model there is no guarantee that the ANN learning algorithm converges and the tuning of fuzzy inference system will be successful but the success of evolutionary search procedures for optimization of fuzzy inference system is well proven and established in many application areas. Abraham (2002) investigated the improvement of the optimization of fuzzy inference systems by using a meta-heuristic approach combining neural network learning and evolutionary computation.

Researchers are contributing in developing to merge all three of ANN, FL and GA approaches but the additional complexity creates new layers of problems such as increased number of parameters that need to be tuned and an increase in the time needed to learn rules (Shapiro, 2002). These types of issues can be resolved by adopting more efficient searches such as Gaussian MFs, since they have only two parameters, rather than three (triangular) or four (trapezoidal), can be used to develop more efficient search strings (Shapiro, 2002). More dynamic routines such as dynamic crossover and mutation probability rates can also be incorporated. Shapiro (2002) suggested implementing simultaneous optimization because the augmented design stages may not be independent or if the new design involves a restricted search space, it may be more likely to result in partial or sub-optimal solutions.

2.1.16 Type-2 Fuzzy Logic Model

Zadeh (1975) introduced the concept of a type-2 fuzzy set which is characterized by a fuzzy membership function, that is, the membership grade for each element of this set is a fuzzy set in [0,1]. The grades of membership of Type-2 fuzzy sets are themselves fuzzy which appeals the researchers because the grades of membership can never be obtained precisely in practical situations (Dubois and Prade, 1982). The membership grade for each element of a type-1 fuzzy set is a crisp number in [0,1]. This membership concept addresses the limitations of ordinary sets in the situations where the membership of an element cannot be determined in a set as 0 or 1. Type-2 fuzzy sets come in the scene when there is uncertainty to exactly determine the membership grade as a crisp number in [0,1] (Karnik and Mendel, 1988). Higher fuzzy set better represents the uncertainty but no finite-type fuzzy set can represent it completely.

Type-2 FLS has been successfully applied in wide range of areas including classification of video streams (Liang and Mendel, 2000a), co-channel interference elimination from communication channels (Liang and Mendel, 2000b), connection admission control (Liang et al., 2000), control of mobile robots (Wu, 1996), decision making (Chaneau et al., 1997; Yager, 1980), solving fuzzy relation equations (Wagenknecht and Hartmann, 1988), survey processing (Karnik and Mendel, 1999 Seoul), time-series forecasting (Karnik, N, N. and J. M. Mendel. 1999), and preprocessing of data (John et al., 1998).

The current research is still limited to type-2 fuzzy sets because the complexity of fuzzy logic system increases rapidly with increasing types. But the fundamentals of fuzzy logic system are same for type-1 to type-n fuzzy sets (Lee and Lee, 2002).

2.1.17 Wavelet Neural Network

Wavelet analysis is a frontier mathematical method which has a good time-frequency local performance and has obvious advantages in describing the non-linear, non-stationary signals (Yang and Zhu, 2006). Zhang and Benveniste (1992) first mentioned

Wavelets networks in the context of non-parametric regression of functions. Zhang and Benveniste (1992) realized that wavelet networks inherit the properties of wavelet decomposition and focused on their universal approximation property, the availability of convergence rates and the explicit link between the network coefficients and the wavelet transform. WN models combine the strengths of discrete wavelet transform and neural network processing and thus have been successfully applied to forecasting and function approximations (Xie and Zhang, 2006). In wavelet networks, the radial basis functions of RBF-networks are replaced by wavelets and during the training phase, the network weights as well as the degrees of freedom (position, scale, orientation) of the wavelet functions are optimized. The nonlinear time series prediction models of Kardanpour et al. (2005); Cui et al. (2005); and Chen et al. (2006) showed that WNmodels were better than either BPNN or RBFNN models produced better prediction results.

Xie and Zhang (2006) proposed two WN models for short-term traffic volume forecasting based on different mother wavelets using Levenberg-Marquardt training algorithm which ensures better efficiency than the other algorithms based on gradient descent. The WN models are better predictor of accuracy, stability, and adaptability compared to BPNN and RBFNN models depending on the error measures including mean absolute percentage error (MAPE) and variance of absolute percentage error (VAPE).

Hai-shuang et al. (2008) proposed a fuzzy wavelet neural networks which takes wavelet function as fuzzy membership function, uses neural networks to realize fuzzy reasoning, and completes the estimate of next cyclical traffic flow. The proposed study used hierarchical genetic algorithm to optimize the network structure and parameters. The network performed better than the conventional BP network in convergence and forecasting accuracy.

2.1.18 Multitask Neural Network

The traditional neural network model for traffic flow forecasting is intended to learn single task at a time (William, 1999) which neglects the potential and rich information resources hidden in other related tasks. MTL is a form of inductive transfer which uses domain specific information which is included in the training signals of extra tasks to improve generalization performance of neural networks (Caruana, 1997). When it is required to solve a complicated problem, it can be split into a number of small and independent sub-problems to learn but it may ignore a potentially rich source of information contained in the training signals of other tasks drawn from the same domain .

Jin and Sun (2008) proposed multitask learning (MTL) based neural networks for traffic flow forecasting which addressed the limitations of single task learning (STL) models which do not take advantage of the information provided by related tasks. Traditional forecasting methods are generally to anticipate with STL. The proposed MTL synchronal to train more than one task and can take full advantage of training information in the extra tasks to improve the generalization of the network which makes the net have higher forecasting accuracy. In their experiments, they found out that the forecasted results of traffic flows are closer to the true value when multitask learning is used in the network.

2.1.19 Group Method Data Handling Based Model

Nearly 30 years of research in the field of neural network and advanced statistical methods contributed in the evolution of abductory induction mechanism (AIM) (Barron et al., 1984). It is a GMDH algorithm which automatically synthesizes abductive networks from a database of inputs and outputs having complex and nonlinear relationships (AbTech Corporation, 1990). AIM provides an environment to synthesize, analyze, and encode abductive networks for complex decision, prediction, control, and classification problems.

Ivakhnenko and Ivakhnenko (1995) listed the basic problems solvable by GMDH, which include identification of physical laws, approximation of multidimensional processes, short-term stepwise forecasting of processes and events, long-term stepwise forecasting, extrapolation of physical fields, clustering of data samples and search for a physical clustering that corresponds to the physical model of an object, pattern recognition in the case of continuous or discrete variables, diagnostic recognition using probabilistic sorting algorithms, self-organization of multilayered neural nets, normative vector of forecasting processes, and process forecasting without models using analogue complexing.

2.1.20 Ensemble Model

Ensemble models which are also known as committee machines refers to the procedures employed to train multiple learning machines and combine their outputs based on the principle that the committee decision, with individual predictions combined appropriately, should have better overall accuracy, on average, than any individual committee member (Brown, 2010). Both empirical and theoretical studies showed that ensemble models very often ensure higher accuracy than single models for predicting real-valued numbers, class labels, posterior probabilities, rankings, and clustering. Their decisions can be combined by many methods such as averaging, voting, probabilistic methods and the majority of ensemble learning methods are applicable across broad classes of model types and learning tasks (Brown, 2010). Ensemble models are mainly developed using heuristics and learning-theoretic principle based algorithms. Brown et al. (2005) and Polikar (2006) reviewed different available algorithms for ensemble models.

In real-world situations, every model has limitations and will make errors. In this situation, ensemble model aims to manage their strengths and weaknesses, leading to the best possible decision being taken overall. If it can be understood precisely why, when, and how particular ensemble methods can be applied successfully, then a significant progress has made toward a powerful new tool having the ability to automatically exploit the strengths and weaknesses of different learning systems and it will cause huge impact across many fields of study (Brown, 2010). In order to improve the performance of classification and regression problems ensemble of neural networks have been frequently employed (Reilly et al., 1987; Neal, 1992, Jacobs et al., 1991; Hansen and Salamon, 1990).

Stathopoulos et al. (2008) suggested a new AI-based approach which addressed the problem of the accuracy of short-term traffic flow forecasting in the complex case of signalized arterial networks. In their approach, a fuzzy rule-based system (FRBS) is augmented with an appropriate metaheuristic (direct search) technique to automate the tuning of the system parameters within an online adaptive rolling horizon framework and the proposed hybrid FRBS is used to nonlinearly combine traffic flow forecasts resulting from an online adaptive Kalman filter (KF) and an artificial neural network (ANN) model. The empirical results showed that their proposed approach perform considerably better than individual traffic predictors.

2.1.21 Summary of Section 2.1

The evolution of traffic flow in time and space considering complex settings such as signalized streets with closely spaced signals and uncontrolled mid-block demand is not clearly understood to develop accurate mathematical model. Due to the dynamically evolving nature of traffic, the parametric statistical approaches/methods can model multivariate and multi steps-ahead forecasting relationships but with the cost of reduced accuracy and adaptability, and increased complexity. The limitations of those models inspired the researchers to develop AI based models for traffic flow prediction mainly using ANN. It is capable of learning any smooth nonlinear mapping using numerical data. But it doesn't provide any explicit representation of knowledge about the system. On the other hand, the fuzzy logic based model allows accurate representation of a given system although it can't tackle knowledge stored in the form of numerical data. In order to get benefitted from both the models different methodologies are implemented to combine them for solving different kinds of prediction problems. There are also a few other models such as support vector regression, abductive networks, wavelet based NN etc. In some cases optimization techniques are also adopted to facilitate the hybrid modeling. Many of those proposed models are not investigated in the area of traffic flow prediction specifically for intersection traffic flow prediction. Generally, the research on traffic flow prediction is limited to short-term duration in order to serve the adaptive traffic signal controllers.

2.2 Literature on TOD Breakpoints

TOD is defined as a mode in which two or more timing patterns are preset to match the traffic demands typically experienced at those times (Wang et al, 2005). In United States, the most widely used method for timing plan selection and implementation is TOD, where a pre-set plan is automatically used for a particular time interval (Gordon et al, 1996). The lack of reliance on detectors is one of primary reasons for the popularity of TOD mode (Wang et al, 2005). In TOD mode, a day is selected into a number of intervals such as the AM peak period, the PM peak period, nighttime in which the traffic volume patterns or demands are relatively stable in each segment (Wang et al, 2005). Although the increased number of intervals will decrease the variation of the demand within each interval and the corresponding optimum traffic signal plan will better serve the traffic, the frequent changes in signal timings may be disruptive to the road users (see Hua and Faghri, 1993).

2.2.1 Statistical Approach

A few statistical and heuristic methods are recently proposed for TOD breakpoints by investigating archived traffic data (Park et al, 2004). The adopted statistical approaches in the literature include both hierarchical and non-hierarchical clustering techniques for finding TOD breakpoints.

Wang et al (2005) applied K-means clustering method to traffic volume data to determine TOD breakpoints for traffic signal timing plans with the assumption that the analyst specifies the number of clusters prior to engaging the algorithm. However, smoothing may be needed to avoid inoperable number of TOD breakpoints which results due the noisy allocations of volumes to clusters. The algorithm was applied to a small case study involving a two-intersection corridor and just less than four hours of volume data.

Smith et al (2001) used data mining tool namely hierarchical clustering algorithm which was intended to support the maintenance of traffic signal systems that operate in the TOD mode by indentifying when traffic conditions have changed significantly. They conducted a case study to “demonstrate that accurate classification models can be developed by using archived data to map between a set of traffic conditions and the associated TOD interval or timing plan for which the conditions are best suited” (Smith et al, 2001). Their approach calls for many transition periods for the timing plans which were even greater than the existing TOD breakpoints and two of the clusters were only thirty minutes in duration. The clusters having very short duration are basically infeasible TOD intervals which are termed as “unclean” clusters by Park et al (2003a). According to them these are inevitable since statistical clustering algorithm do not consider time variable.

Park and Lee (2008) present the development and evaluation of a procedure for determining optimal break points for TOD based coordinated actuated traffic signal operations which uses a feature vector of optimal cycle length per time interval instead of traffic volume itself. Initial breakpoints determined by the feature vector are used in greedy search algorithm which requires less number of evaluations in the search compared to an exhaustive search or other common heuristic search methods such as GA and finds out the optimal breakpoints. Greedy algorithms build up a solution piece by piece, always choosing the next piece that offers the most obvious and immediate benefit without worrying about the effects of these decisions may have in future. This sort of algorithm usually does not operate exhaustively on all the data. The results for a hypothetical network consisting of four signalized intersections reveal better performance of the coordinated actuated signal control when the proposed procedure is adopted.

2.2.2 *AI and Other Approaches*

Park et al (2003a) enhanced data driven methodology to overcome the drawback by adopting a GA based clustering which is evaluated through the performance of actual timing plan under the simulation software “SimTraffic”. But the process of TOD breakpoints optimization does not consider the signal timing plan. Park et al (2004) proposed a GA based method which implements two-stage optimizations: outer loop for TOD breakpoints and inner loop for timing plans of corresponding intervals on a network consisting of three coordinated actuated signalized intersections. The six number of breakpoints outperformed the other considered numbers of breakpoints in the simulated environment.

Abbas (2006) introduces a multi-objective evolutionary algorithm with Degree of Detachment (DOD) which is successful in finding optimal selection of timing plans as well as the optimal TOD schedule. The evolutionary algorithm does not account for the need of engineers to have minimum number of transitions between timing plans during the day (Abbas et al., 2005). Abbas et al. (2005) emphasized on the continuation of a timing plan and defined the DOD metric for the purpose of clustering traffic patterns. The DOD measures the degree, by which the traffic pattern at the time period is detached from adjacent periods in term of its assigned timing plan. The DOD is used to avoid zigzag changes in timing plans and it is introduced as a performance measure of scheduling continuity. High DOD translates into frequent changes in timing Plans and a zero DOD translates into a one timing plan applied throughout the day.

Hua and Faghri (1993) presented a dynamic programming (DP) method for identify TOD breakpoints which is based on finding an optimal set of control parameters, including the cycle length, splits and offset, for a given time interval. They also presented an ANN approach which is based on the searches for traffic patterns that are similar to "mother patters" in order to find out appropriate TOD breakpoints. The "mother pattern" is a pre-selected traffic pattern but the authors didn't provide any guidance on its selection. Generally, the ANN approach is computationally less intensive compared to DP.

Smith et al (2002) proposed a new methodology for developing timing plan that automates the identification of TOD intervals using a high-resolution definition of system state. They claimed that the proposed methodology supports the development of a TOD system that provides benefits when considering performance measures such as delay, when compared to currently used techniques.

2.2.3 Optimization of Transition Period

According to Mussa and Selekwa (2003), one of the downsides of the TOD timing technique is the handling of coordination parameters such as cycle length, phase split, and offset during the transition between two plans. Their methodology optimizes traffic flow during the transition period which is based on dynamic quadratic optimization, achieves synchronization of coordination parameters. The methodology of optimizing transition period has the potential of reducing queue delay, particularly on minor street approaches (Mussa and Selekwa, 2003).

There are practical and theoretical methods for optimizing traffic flow during transition period resulting from applying different plans for different intervals of TOD. Hongqiang et al (2005) practically improved the optimization during transition period by optimizing first the offset and subsequently the cycle length. The optimization of offset minimizes the disturbances caused by the transition. The optimization of the cycle length minimizes transition duration and provides sufficient capacity.

2.2.4 Cluster Validation

Cluster validation is important in most applications including the problem of TOD breakpoints. There are a lot of approaches to evaluate the performance of the cluster. The Dunn's Validity Index is based on the idea of identifying the cluster sets that are compact and well separated (Dunn, 1974). Davies-Bouldin Validity Index is a function of the ratio of the sum of within-cluster scatter to between-cluster separation (Davies and Bouldin, 1979). The Isolation Index technique is based on the assertion that neighboring instances in feature space often occur in the same natural cluster (Pauwels and Frederix, 1999). Hubert and Schultz (1976), Goodman and Kruskal (1954), Jaccard (1912), Rand (1971), Topchy et al (2003) provided different approaches for cluster validation. The interested readers can consult the provided references. The quality criterion called silhouette is used in order to find out correct number of clusters (Lamrous and Taileb, 2006; and Wong and Woon, 2008). Rousseeuw (1987) introduced 'silhouette' as a general graphical aid for interpretation and validation of cluster analysis.

2.2.5 Application of Cluster Analysis in Traffic Engineering

Prassas et al. (1996) suggested cluster analysis as a solution to traffic engineering data based on the premise that there are a finite number of distinct "states" or conditions in which the system may rest, with significant truly random deviations around each such equilibrium. The premise is different than that of deterministic modeling and regression analysis. They applied cluster analysis tool to a set of traffic engineering data (left-turn factors in shared lanes). They found out that cluster analysis is a powerful exploratory technique and helped identifying several distinct modalities within the data.

Niemeier et al. (2002) presented a cost-efficient data sampling approach to estimate mobile emissions. They used hierarchical cluster analysis on the correlation (similarity) matrices to statistically identify groups or clusters of count locations with similar hourly profiles. Then a single count within each cluster was chosen to represent the hourly diurnal profile of count proportions. The proposed approach helped in determining the key counts needed to both minimize data costs as well as provide sufficient and robust information for subsequent modeling.

Weijermars and Berkum (2005) used Ward's hierarchical clustering procedure to determine historical traffic patterns. It was observed that working days are easier to classify into distinctive, recurrent traffic patterns compared to non-working days. On the basis of the resultant classification of working days they concluded that four types of working days, can be distinguished: (1) Mondays, (2) core week days, (3) Fridays and (4) days within vacation periods.

Chen et al. (2006) proposed a traffic data analysis of ITS in which Kernel Principal Component Analysis (KPCA) was used to reduce data dimensionality and extract features from them, then Self-Organizing Map (SOM) was applied in the unsupervised clustering of links. They demonstrated that Kernel PCA outperforms traditional PCA in data compression for it take into account the non-linear patterns in traffic data, and that SOM can reveal underlying relations between different links. The outcome of the proposed approach might serve the research on traffic parameter forecasting and flow control.

Lee et al. (2004) used data mining technique in investigating an incident situation. They demonstrated that it can divide incident data into clusters according to flow, density and time. It enables the user to demark the impact area of the incident temporally and spatially. The information represented via visualization aid can assist Traffic Management Center (TMC)/traffic engineers to take necessary actions at the occurrence of an incident and to identify any hidden relationship among traffic data.

2.2.6 Review of Traffic Simulation Models

Simulation is the abstraction of real world conditions by developing computer model and running them through time (Ni, 2001). Five driving forces contribute significantly in the advancement of traffic simulation: the advances in traffic theory; the continuing improvement in computer hardware and software technology; the development of the general information infrastructure; and the society's demand for more detailed analysis of the consequences of traffic measures and plans (Ni, 2001). Traffic systems are complex systems because of human interactions and man-machine interactions. Due to the complexity of this type of system, simulation is required to test, evaluate and demonstrate a proposed course of action before implementation. The descriptions of the traffic simulation models including both macroscopic and microscopic are summarized in Table 2-1. In the subsequent paragraphs the main features of a few simulation models are described.

AIMSUN integrates in a single software application three types of transport models. These are traffic assignment models, a mesoscopic simulator, and a microsimulator (TSS, 2008). The microscopic model is developed based on car following, lane changing, and gap acceptance algorithms like other software such as CORSIM. However, the new mesocopic simulator in AIMSUN 6 provides an additional option to practitioners to model dynamic aspects of very large networks and removes most of the calibration burden when compared to a micro-simulator (TSS, 2008). The software's home page provides demo software along with the many relevant resources.

SimTraffic is a microscopic simulation package which uses the SYNCHRO program to model street networks. It was initially developed to model the arterial signal system timings. It can simulate surface street networks, freeways, weaving sections, pre-timed and actuated traffic signals, stop-controlled intersections, roundabouts, transit operations, pedestrians, etc. (Trafficware, 2008).

VISSIM is based on a traffic flow model which is discrete, stochastic, and time step based microscopic model. The model considers driver-vehicle-units as single entities and contains a psycho-physical car following model for longitudinal vehicle movement and a rule-based algorithm for lateral movements. The model is developed based on the research of Wiedemann (Wiedemann, 1974 and 1991). The simulation package VISSIM consists internally of two different programs which include the traffic simulator**,** a microscopic traffic flow simulation model and signal state generator. The latter is a signal control software polling detector which compiles information from the traffic simulator on a discrete time step basis. Fellendorf (1994) described the system architecture of VISSIM concentrating on its abilities as a simulation model for signal control. The simulation systems of VISSIM consist of traffic flow model and signal control model. VISSIM sends detector values to the signal control program in every second and the signal control uses the detector values to decide the current signal aspects (Fellendorf, 1994).

ACTSIM is a dynamic micro-simulation model which simulates each vehicle independently, uses the distribution of population behavior, models changes in density, peaking in demand, curbside parking and crosswalks, and provides a continual picture of the network. Statistical data gathered over a period provides an average view as well. Individual vehicle parameters including vehicle speed, vehicle size, desired maximum speed, destination location, dwell time, and gap acceptance are assigned by random variants derived from vehicle mode characteristics. ACTSIM consists of a car following model, lane changing model, parking model, pedestrian crossing model, and passenger pickup/drop off model (Ni, 2001).

CORSIM is a microscopic, stochastic, link-node and periodic-scan based traffic simulation program designed for the analysis of freeways, urban streets, and corridors or networks. The combination of arterial (TRAF-NETSIM) and freeway (FRESIM) simulation models make CORSIM one of the analysis tools available to traffic engineers that allow all of the individual components of the arterial and freeway system to be analyzed and simulated as a complete system (Prevedouros and Wang, 1999). CORSIM stochastically determines the specific properties of each vehicle such as vehicle length, driver aggressiveness, acceleration rate, minimum acceptable gap, maximum free speed, and others. The car-following and lane-changing logic to simulate vehicle movements are done in CORSIM on a second-by-second basis.

Quadstone PARAMICS is a suite of software models for microscopic traffic simulation which allows a unified approach to traffic modeling encompassing the whole spectrum of network sizes starting from single junctions up to national networks. It models the emerging ITS infrastructures. Programmer, which is a software development kit (SDK) for the PARAMICS suite, can be used for research of all aspects of ITS, real time connectivity and control, connectivity to real world hardware and software systems, and advanced or customized model behaviors (Quadstone, 2008). The software's home page provides demo software along with the source information of many relevant resources.

TRANSYT is an off-line macroscopic deterministic simulation and optimization model that simulates traffic as cycle flow profiles (CFP), traces the flow of CFP from link to link throughout the network, and makes systematic changes to the offset, phase split, and cycle length of the traffic signals. It also simulates the associated traffic conditions to estimate a corresponding performance index (PI). This PI is composed of vehicle delay and number of vehicle stops. The simulation module within the TRANSYT model evaluates the objective function that is to be minimized (Rakha and Van Aerde, 1996).

Table 2-1: A few different types of simulation models and their main features and capabilities.

Name	**Characteristics**	**Main Features/Capabilities**
CORSIM	Microscopic	Surface streets, freeways, actuated signals, weaving sections, incidents, variable message signs, 2-D animation.
SimTraffic	Microscopic	Surface streets, actuated signals, pedestrians, roundabouts, 3-D animation.
GETRAM	Microscopic, distributed computing technique	Surface streets, freeways, actuated signals, dynamic traffic assignment, variable message signs, 3-D animation, telematics.
VISSIM	Microscopic	Surface streets, freeways, ramp metering, pedestrians, transit operations, 3-D animation.
PARAMICS	Microscopic, Distributed computing technique	Surface streets, freeways, transit operations, 3-D animation, roundabouts, congested networks.
INTEGRATION	Mesoscopic	Surface streets, freeways, traffic assignment, intelligent transportation system, toll plaza, vehicle emissions, HOV.
DynaMIT	Mesoscopic, real time computer system	Operation of Advanced Traveler Information Systems (ATIS) and Advanced Traffic Management Systems (ATMS), dynamic estimation of network state, a variety of real time scenarios, simulation of each trip.
MITSIMLab	Microscopic	ATIS and ATMS.
CORFLO	Macroscopic	Surface streets, freeways.
SATURN	Microscopic	Individual junctions, traffic assignment.
Micmac	Hybrid	SITRA B+ (microscopic model) and SIMRES (macroscopic model) are coupled, and the synchronization of the models is sequential.

Name	Characteristics	Main Features/Capabilities
Hystra	Hybrid	Macroscopic and microscopic models are combined, both models are based on the (Lighthill-Whitham-Richards) LWR traffic flow theory.
FRESIM	Microscopic	Freeway blockage incidents, variations in grade, lane additions, lane drop anywhere in the freeway and auxiliary lanes.
KRONOS	Macroscopic	Freeway lane changing, merging, diverging, and weaving, the simultaneous development of queues and propagation of congestion on both the freeway and its ramps.
KWaves	Macroscopic, discrete, deterministic	Freeways, throughput, bottlenecks, queues, ramp metering, incident management.

(Source: Ratrout and Rahman, 2009)

2.2.7 Summary of Section 2.2

In developing countries and even in many developed countries the most widely used method for timing plan selection and implementation is TOD. But there is very little research found in the area of determining TOD breakpoints. The statistical clustering based approach typically fails to ensure clean clusters. There are also a few evolutionary algorithm based approaches to solve the problem of unclean clusters. In some cases, MOEs are considered to determine optimum TOD breakpoints. An automated procedure to determine TOD breakpoints can solve the subjective approach prevailing in the developing countries.

CHAPTER 3 METHODOLOGY AND DESCRIPTION OF THE RESEARCH

This methodology provides a framework to obtain traffic flow prediction of an intersection and to determine optimum TOD breakpoints. The proposed procedural framework can be divided into three main stages. In the first stage, a number of prediction models are investigated to predict the Approach traffic flows of an intersection using its and other neighboring intersections' traffic flows. In the second stage, the predicted Approach traffic flows of the intersection are used to determine optimum TOD breakpoints. In the third stage, the obtained TOD breakpoints using observed traffic flows and predicted traffic flows are investigated through microscopic simulation model. The methodology to develop hybrid AI model for freeway traffic flow using local data of Saudi Arabia is reported in the Chapter 6.

3.1 Traffic Flow Prediction

Due to the nonlinearity and uncertainty of intersection traffic flow, this study attempts to use AI based models such as GMDH, ANFIS, type-2 fuzzy logic and ANN model. GMDH model is deterministic and self-organizing in nature and it requires minimum intervention from the user. These features make it useful for the transportation engineering practitioners. ANFIS model exploits the individual capabilities of ANN and fuzzy logic model. It provides insights about the problem and it has the learning capability like ANN. In the process of ANFIS model building it is not required to provide the rule-base explicitly. Type-2 fuzzy model is capable to handle the uncertainty associated with measurement and with the membership functions which are predefined in type-1 fuzzy model. These inherent characteristics of Type-2 fuzzy model are not explored extensively in developing traffic flow prediction model. The variations of those mentioned models which are used in this study are not explored earlier in traffic flow prediction problem.

ANN model is also investigated in this study to determine the comparative performance of the selected models. There are very few or no applications of neuro-fuzzy and type-2 fuzzy logic models available in the field of traffic flow prediction and only a few variations of those models are explored. It is also note-worthy to mention that the GMDH model is not explored in the field of traffic flow prediction as per the knowledge of the author.

3.1.1 Input Selection

The model building process of any AI model starts with the proper selection of input data. The GMDH model which is explored in this study (AIM model) has an embedded procedure to select the most important input variable. The model is self organizing in nature and it doesn't require much intervention of the user. The user can provide some optional parameters which are related to complexity of the network and number of input for the first layer. The input variables for neuro-fuzzy model are mainly selected depending on the principle stated by Jang (1996). Generally, the input variables which depend on other input variables are not considered in building the model. Then different input variables are selected to predict the approach traffic flow of the desired intersection. In this case, the performances of only a few epochs were considered to select the input variables. The nuro-fuzzy model specifically, the ANFIS uses a hybrid learning method that combines gradient descent and the least squares method. As the least-squares method is the major driving force of ANFIS, it can usually generate satisfactory results just after the first epoch of training. The input variables for type-2 FLS are selected in trial and error basis.

The ANN models were also developed for comparative studies. When the ANN model is compared with type-2 FLS then the input variables are same as that of the type-2 FLS. This concept is followed when the ANN models are compared with GMDH and ANFIS models.

3.1.2 Input Processing

The efficient, accurate and robust AI model building heavily relies on the appropriate preprocessing of the input variables. The GMDH based AIM model preprocessed the input variables through mean-sigma normalization. The input variables of the type-2 FLS were not preprocessed as the output seemed to be less sensitive with preprocessing. The input variables of ANFIS models were preprocessed to make the values of input variables within the range of -1 to +1. The input variables of ANN models were transformed in such a way that the mean and standard deviation of each variable became zero and one respectively.

3.1.3 Model Building

The model building approach differs depending on the type of model. In this sub-section brief description of model building approach for the selected model are discussed. The details of the individual model building approach are discussed in Chapter 4.

AIM model depends on a few important parameters such as complexity penalty multiplier (CPM), limiting number of layers, and number of inputs. The effects of those parameters on the error measures are investigated to develop better performing networks.

The direct implementation of ANFIS causes huge number of rule generation. In order to solve this problem, the initial fuzzy inference systems (FISs) are developed by using fuzzy C-means (FCM) clustering and subtractive clustering methods. In this study, only generalized bell shaped membership functions (MFs) for inputs were tried because the choice of MFs of ANFIS architecture doesn't have significant impact on the model performance.

In order to develop type-2 fuzzy logic model, all the fuzzy input sets and the random center of consequents are used to build the initial rules which are equal to the sample size. The Singular Value Decomposition (SVD) technique is applied to reduce the number of rules and determine the most important rules to produce desired output. In this study, all the proposed non-singleton type-2 models assume Gaussian primary membership functions with uncertain mean and interval secondary membership functions, product implication and t-norm, center-of-sets type reduction, and defuzzification obtained by the centroid of the type-reduced set.

3.1.4 Model Evaluation

There are many error measurements which are used to evaluate the performance of prediction model. In this study, a few error measures are considered which include Root Mean Square Error (RMSE), Mean Absolute Error (MAE), R-square, Average Error, Standard Deviation of Error. In this study, Regression Error Characteristic (REC) analysis is also performed to provide greater insights about the error of the proposed models.

In order to determine acceptable error measures a few relevant studies are consulted. Jin and Sun (2008) developed NN through multi-task learning for traffic flow prediction of a few road links of an urban area using 15-min traffic counts. The prediction period was 3 days. The reported RMSE was 88.02 vehicles/hour for a link and the peak hour volume of that link was around 1400 vehicles/hour. Afandizadeh and Kianfar (2009) reported the error of their hybrid neuro-genetic model for predicting rural highway traffic flow. The Mean Square Error (MSE)s were 63.25 and 261.77 vehicles for 5 minute and 15 minute interval respectively. Afandizadeh and Kianfar (2009) didn't provide any information about the typical traffic flow characteristics of the considered network. Yu and Zhang (2004) proposed a switching ARIMA model and reported the errors of the model. The mean root square error (MRSE) and the mean absolute relative error (MARE) of their model was 105.2093 vehicle/hr and 9.95%. The corresponding peak hour traffic flow of the considered two days was around 1900 vehicles/hour.

In this study, it is assumed that RMSE less than 4 vehicles/15 min for Approach 5 and 6, and RMSE less than 13 vehicles/15 min for Approach 7 and 8 are acceptable. These assumptions are made after carefully reviewing the previous reported error measures for intersection traffic flow prediction, and the probably consequences or effects in terms of MOEs.

3.2 TOD Breakpoints Determination

The determination of TOD breakpoints depends on the selection of features. The earlier hierarchical, and non-hierarchical clustering technique such as K-means based approach to find out optimal TOD breakpoints suffered from “outliers” of clusters although they considered a list of features such as volume, occupancy etc. In this study, Z-score of all approaches of the Intersection A and time variable expressed in hour are considered as input features for implementing clustering technique. The use of Z-score instead of the traffic counts ultimately provides greater relative importance to the time variable in determining the cluster. In order to investigate the effect of features in K-means algorithm the 15-min traffic counts were clustered to determine number of transitions. The traffic counts along with time variable were also used to determine the number of transitions.

In order to get consistent and better results in K-means algorithm the initial cluster centers were obtained through subtractive clustering method. This method determines the initial cluster centers and the number of clusters considering the four approach volumes as the input. Based on the experiments, the radius value of this algorithm is to be fixed. The obtained cluster centers are provided as the initial solutions for the K-means technique to determine TOD breakpoints. K-means considers time variable and Z-score of the approach volumes. Finally, K-means determines the TOD break points for observed and predicted data.

3.3 Evaluation through Simulation

In this section, it is attempted to elaborate and demonstrate the methodology of evaluating the performance of TOD breakpoints through microscopic simulation such as SimTraffic. This software is chosen due to the wide acceptance of this model in Saudi Arabia. Ratrout and Olba (2009) commented that SimTraffic is one of the models which is being used extensively by local engineering consultancy offices. Due to the userfriendly nature of this model it became popular among the practitioners in Saudi Arabia (Ratrout and Rahman, 2009).

In the proposed methodology, it is assumed that the effects of transition will be limited to the duration of one hour. For example, an analysis reveals two TOD breakpoints at 07:00 and 13:00 which corresponds to time T-1 and T-2 in Figure 3-1. In this case, three different optimum traffic signal plans namely Timing Plan I, Timing Plan II, and Timing Plan III will be used between 00:00 to 07:00, 07:00 to13:00, and 13:00 to 00:00, respectively as shown in Figure 3-1. Timing Plan I, II, and III are the optimum timing plans which are developed based on the design traffic flows D-1, D-2, and D-3 respectively which are graphically shown in Figure 3-1.

In order to find out the effects of a breakpoint at a certain time such as 07:00 (T-1 in Figure 3-1), the actual traffic counts from 06:30 to 06:45, and from 06:45 to 07:00 was simulated under the Timing Plan I. At the same time, the actual traffic counts from 07:00 to 07:15, and from 07:15 to 07:30 was simulated under the Timing Plan II. The time duration from 06:30 to 07:30 is represented by S-2 in Figure 3-1. The remaining periods of each TOD such as S-1 and S-3 was simulated by using observed hourly traffic counts under Timing Plan 1 and Timing Plan 2 consecutively. The observed hourly traffic counts are supposed to be used due to the limited number of intervals (only 19 intervals) which can be handled by SimTraffic. By adding the values of MOEs for 24 hours the performance of TOD breakpoints can be evaluated. The number of clusters which will ensure minimum overall cost with respect to the predefined MOEs, will govern the selection of TOD. In order to make the obtained MOEs more reliable the study suggests using multiple microscopic simulation runs in SimTraffic by adopting random seeds. The readers can consult the manual to understand the methodology adopted in SimTraffic (Husch and Albeck, 2006).

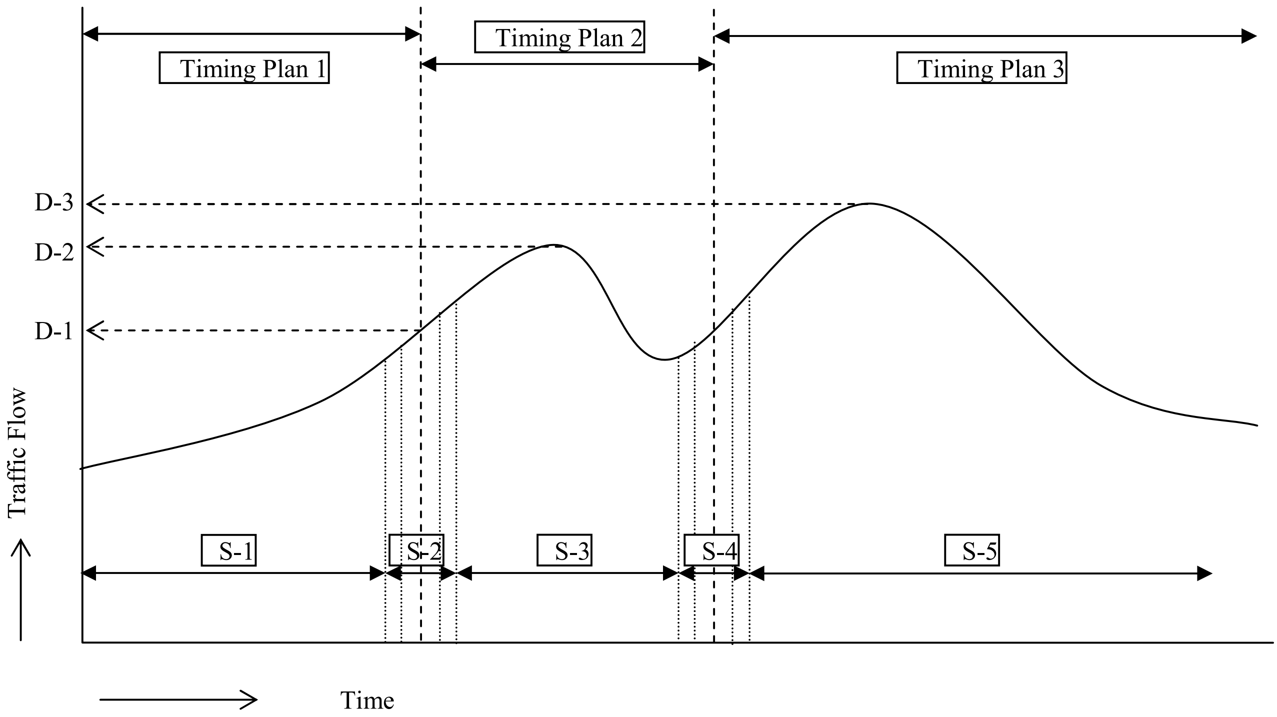


Figure 3-1: A Graphical presentation of TOD breakpoints and relevant information.

3.4 Summary of Chapter 3

This chapter provides a methodology to obtain traffic flow prediction of an intersection and to determine optimum TOD breakpoints. The proposed procedural framework can be divided into three main stages. In the first stage, a number of prediction models are investigated to predict the approach traffic flows of an intersection using its and other neighboring intersections' traffic flows. In the second stage, the predicted approach traffic flows of the intersection are used to determine optimum TOD breakpoints. In the third stage, the obtained TOD breakpoints using observed traffic flows and predicted traffic flows are investigated through microscopic simulation model.

The model development process of the individual model depends on input selection, pre-processing, model building procedure. The input selection, pre-preprocessing and model building process of the GMDH based abductive network are self-organizing in nature and it doesn't require much intervention of the users. The input data were selected through standard method for ANFIS, type-2 and ANN models. For ANFIS models, the initial fuzzy inference systems (FISs) are developed by using FCM clustering and subtractive clustering methods. In order to develop type-2 fuzzy logic model, SVD technique is applied to reduce the number of rules and determine the most important rules to produce desired output.

TOD breakpoints are determined through subtractive clustering based K-means method. The initial cluster centers of K-means algorithm are provided by subtractive clustering method. This approach is adopted for both predicted and observed traffic data. Then the performances of the TOD breakpoints are evaluated through microscopic simulation model with respect to MOEs. The hourly data are used to evaluate the each TOD and 15 min traffic counts are used to evaluate the transition points. The traffic signal timing plan changes at the location of transition point.

CHAPTER 4 TRAFFIC FLOW PREDICTION

4.1 GMDH BASED ABDUCTIVE NETWORK MODEL

Typically, GMDH algorithm is self-organizing in nature and it doesn't require much intervention from the user to build the model. Due to these features of this algorithm, the author attempts to explore it for predicting traffic flow. This algorithm has been successfully used in many identification and forecasting processes including financial systems, ecological processes, control applications, and diagnostic processes. But this promising AI model has not been explored yet in the literature for traffic flow prediction as per the knowledge of the author.

4.1.1 Concept of GMDH

In modeling complex systems of unstructured areas, problems arise due to the prejudices introduced into the model and the possible consequence that majority of the obtained results are vague, ambiguous, and qualitative in nature (Farlow, 1981). In order to solve the limitations of the structured models, Ivakhnenko (1966) introduced the GMDH algorithm which provides an objective model of a high-order polynomial in the input variables to solve prediction, identification, control synthesis, and other system problems.

Figure 4-1 shows the flowchart of GMDH algorithm. As shown in the figure, this algorithm starts with regression equations of 2 or 3 orders for each pair of input variables and the output. Thus, for "n" number of input variables there will be "n(n-1)/2" higher-order variables which will be used for predicting the output instead of the original n number of input variables. The best performing equations will replace the original input variables. If the regression equations are of 3 orders for two variables, then in the second stage the equations are basically a collection of polynomials of 9 orders.

The algorithm selects the relationships which have good predicting powers within each phase, prevents exponential growth, and limits model complexity. When the prediction performance criteria applied to checking datasets for a new generation starts increasing compared to that of previous generation, then the model basically reached the minimum value of the criteria. This concept of optimality is also verified through experimental works. After determining the optimum generation, the best of the polynomials in that generation is selected. It indicates that the output is now a polynomial of two variables which are not the original inputs.

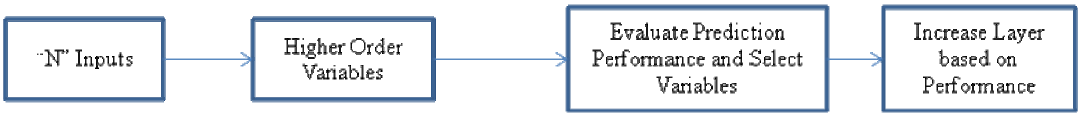


Figure 4-1: The steps followed to implement GMDH algorithm.

The model building process of GMDH algorithm typically involves four steps which are the main features of GMDH theory (Ivakhnenko, 1970): (i) collect relevant set of observations; (ii) divide the datasets into two groups and one of which will be used to estimate the coefficient of the model, and the second datasets will be used in evaluating the performance of prediction; (iii) create a set of elementary functions where an iterative procedure will produce different models; and (iv) apply an external criterion to choose the optimum model. External criterion describes the requirements for the model, for example, minimization of least squares, and it is always calculated with a separate part of data sample that has not been used for estimation of coefficients. In this algorithm, the parameters are estimated based on the training data and there is no specific guideline for selecting the number of training datasets. The checking datasets do not contribute in the parameter estimation. The selection of training and checking datasets is more crucial for a small number of datasets because the different ways of division of datasets may result in very different final models (Mehra, 1977).

The attractive features of GMDH include: (i) determination of the structure of nonlinear systems, (ii) solution of the problem of overfitting by separating training and checking data, and (iii) capability of consideration of multicriteria objective function for validation (Mehra, 1977). The extensive number of GMDH and its ability to model ill-defined problems with reasonable accuracy have proved and strengthened its position as an appropriate nonlinear method for structural identification and prediction tasks (Anastasakis and Mort, 2001).

4.1.2 GMDH Algorithm Based Software Packages

The wide possible applications of GMDH algorithms motivated in developing software packages. The advantages of the software based on inductive approaches include the requirement of minimal a priori information about the system, very fast and effective learning process, model development on very small and noisy data samples, an optimal and cross-validated model, and an analytical model. Ward Systems Group, Inc. developed the GMDH module of “NeuroShell2” tool which uses partial polynomial optimization for network structure construction (Ward Systems Group, Inc., 2009). Muller and Lemke (1999) proposed a self-organizing model known as “KnowledgeMiner”. The model realizes twice-multilayered neuron networks with active neurons, optimizing the structure of every neuron and adaptively synthesizing the network structure and can be used for generation of systems of equations (Lemke and Muller, 1997a). It is designed to support the knowledge extraction on a highly automated level and has been successfully applied in financial systems for prediction tasks or decision systems (Lemke 1997; Muller and Lemke, 1995). The main SKAT module in the “PolyAnalyst” software from “Megaputer Intelligence” uses inductive GMDH-type technique, with sorting of ratio-polynomial models for knowledge discovery and data mining. The Peak Consulting Inc., introduced the “AutoNet” which is an Excel-based application with an open VBA code. The model is a GMDH-type neural network program with a self-designing architecture (Peak Consulting Inc., 2009).

4.1.3 Applications of GMDH

GMDH algorithms have been successfully designed for many identification and forecasting processes including financial systems, ecological processes, control applications, and diagnostic processes. Goleusov and Kondrasheva (1987) investigated the ability of GMDH to extract possible information about the interdependencies between the financial indices of selected countries. Ivakhnenko et al. (1989) concluded that it is more than a simple modeling method. It expands the possibilities of contextual interpretation of the results of economic modeling. Lemke and Muller (1997b), and Water et al. (1997) applied the GMDH algorithm in the prediction of stock prices. GMDH algorithms are also popular in ecological processes which are typically highly nonlinear with noisy data. Krotov and Kozubovskiy (1987) used GMDH algorithm for the prediction of tree-growth rings. Chang and Hwang (1999) applied this algorithm for flood forecasting. GMDH algorithms have also been successfully applied to control applications. Kozubovskiy and Kupriyanov (1987) applied GMDH algorithm in order to calculate the pressure difference at the output of a differential pneumatic bridge which is part of a general pneumatic system. There are many other applications reported in the literature. Abdel-Aal et al. (2009) modeled and forecasted the mean hourly wind speed time series by using a GMDH-based abductive network. Abdel-Aal (2004a) applied GMDH algorithm in forecasting hourly temperature which is important for electrical load forecasting and other applications in industry and agriculture. Abdel-Aal (2004b) implemented GMDH algorithm for modeling and forecasting short-term load which is essential for the profitability of power utilities. Kondo et al. (2005) used GMDH algorithm for medial image recognition.

Ivakhnenko and Ivakhnenko (1995) listed the basic problems solvable by GMDH, which include identification of physical laws, approximation of multidimensional processes, short-term stepwise forecasting of processes and events, long-term stepwise forecasting, extrapolation of physical fields, clustering of data samples and search for a physical clustering that corresponds to the physical model of an object, pattern recognition in the case of continuous or discrete variables, diagnostic recognition using probabilistic sorting algorithms, self-organization of multilayered neural nets, normative vector of forecasting processes, and process forecasting without models using analogue complexing.

4.1.4 Abductive Network

Montgomery and Drake (1991) defined induction as the act or process of reasoning from facts to general principles. Peirce (1955) introduced the term abductory induction as a unique class of induction, which derives abductive principles from facts under uncertainty using numeric functions and measures. It synthesizes abductive principles from empirical observations. If there is a relationship among data, it provides a modeling solution by subdividing the input variables into groups, directing them to individual nodes of a network, summarizing the relationships among each group and feeding these results to the next layer of the network (Montgomery and Drake, 1991). Nearly 30 years of research in the field of neural network and advanced statistical methods contributed in the evolution of AIM (Barron et al., 1984). It is a GMDH algorithm which automatically synthesizes abductive networks from a database of inputs and outputs having complex and nonlinear relationships (AbTech Corporation, 1990). AIM provides an environment to synthesize, analyze, and encode abductive networks for complex decision, prediction, control, and classification problems. The encoded computer subroutines provided by AIM can easily be incorporated into other models as needed.

An abductive network is a network of functional nodes in which the input data is processed through the network to the output variables even allowing feedback loops. This concept has been present in neural and polynomial network research which has been applied in many forms (Montgomery and Drake, 1991). Montgomery (1982) developed an approach to speech recognition which is used to subdivide the problem and reduce the difficulty in determining the polynomial operators.

AIM uses mathematical models to represent complex and uncertain relationships along with polynomial networks representing the underlying process (Buck and Nelson, 1992). Figure 4-2 shows an example of AIM network and its functional elements. AIM uses “Normalizers” and “Unitizers” for the input and output layers, respectively. “Normalizers” transform the input variables through mean-sigma normalization, and “Unitizers” change the range of network outputs to a range with the mean and variance of the output values of the training data. The nodes of a typical feed-forward AIM network can be Singles, Doubles, and Triples which are third order polynomials with one, two or three inputs respectively (Figure 4-2). A Double node is defined as follows:

$$\text{Output} = w_0 + w_1x_1 + w_2x_2 + w_3x_1^2 + w_4x_2^2 + w_5x_1x_2 + w_6x_1^2x_2 + w_7x_1x_2^2 + w_8x_1^3 + w_9x_2^3 \quad \ldots\ldots (1)$$

where x_1 and x_2 are inputs, and w_n (n=1 to n) are the obtained weights after training.

A White node consists of linear weighted sum of all outputs of a previous layer and constant such as:

$$\text{Output} = w_0 + w_1x_1 + w_2x_2 + \ldots + w_nx_n \quad \ldots\ldots (2)$$

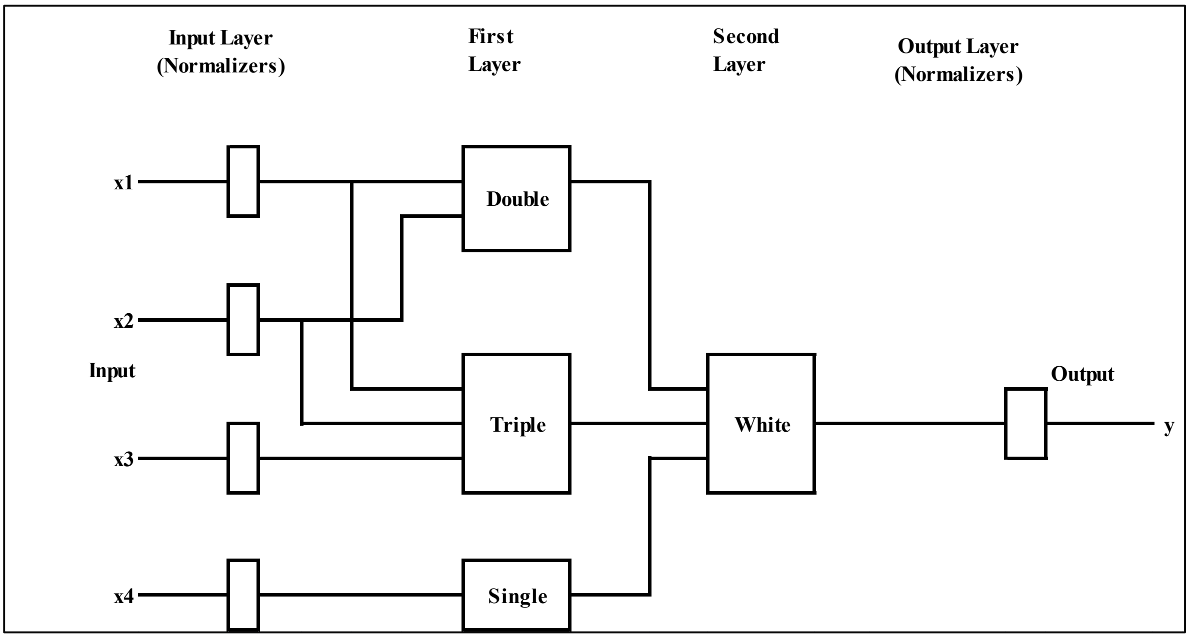


Figure 4-2: An example of AIM network and its functional elements.

AIM simplifies model development and reduces the learning or development time and effort by automatically determining the optimum model characterized by network size, element types, connectivity, and coefficients (Abdel-Aal, 2004a). The selection criteria of the model lead to determine the possible accurate network model without overfitting the training data. It is accomplished by performing a trade-off between model complexity and accuracy (Montgomery and Drake, 1991). Barron (1984) derived the modeling criterion of AIM network which is known as predicted squared error (PSE).

$$\mathrm{PSE} = \mathrm{FSE} + \mathrm{CPM}\frac{2K}{N}s_p^2 \quad \text{...}(3)$$

where PSE is the Predicted Squared Error, FSE is the fitting squared error of the model based on the training data, K is the total number of coefficients used in the model, N is the total number of training data, s_p^2 represents the estimated true unknown model error variance, and CPM (Complexity Penalty Multiplier) is the user-defined penalty multiplier.

The value of the second term in the above equation increases linearly and the first term decreases when the model becomes more complex compared to the size of the training set. For an optimum model size, PSE passes through a minimum, which ensures a balance between accuracy and simplicity (Abdel-Aal, 2004a). The user-defined value of CPM parameter can affect this trade-off. Generally, the larger values of CPM compared to default 1 lead to simpler models with less accuracy, while the lower values of CPM overfit the training data with degraded actual prediction performance (Abdel-Aal et al., 2009).

4.1.5 Model Development

AIM automatically determines the network size, element types, connectivity, and coefficients for the optimum model and it also provides the flexibility to optionally change a few important parameters such as CPM, limiting number of layers, and number of inputs. This study attempted to explicitly investigate the effects of those parameters on the error measures, aiming to develop better performing networks.

The AIM model continues to build a network by increasing the number of layers as long as the resulting network performance continues to improve without exceeding the limit specified by the user. Usually, the typical range of the limiting number of layers is 1 to 9 for many applications, but an increase in the layer limit may increase the time it takes to synthesize a network. In this study, the abductive networks for predicting the four (4) approaches of intersection 15 were investigated for different numbers of layers. The results indicate that the six limiting number of layers performed reasonably well in all the models investigated and increasing the number of layers beyond that value was not improving the model.

An important parameter of the AIM model which can be changed by the user is the size of the first layer. At each layer, the AIM model keeps a list of candidate sub-networks for inclusion in the next layer and the list decreases in size at each subsequent layer. With the increase in the size of first layer, the network synthesis time increases because the AIM model has to test even more network combinations. However, a low value of this setting may risk not finding a successful network topology. A different size of the first layer was investigated in order to find out the typical size which can perform reasonably for all the models. In this study, the abductive models were built with twelve numbers of inputs as maximum size of the first layer unless mentioned otherwise. The model selects the inputs from the available twelve numbers of candidate inputs which include the approach traffic flows of intersections 13, 15, and 17.

PSE takes care of both the fitting squared error of the training data and complexity of the network, and goes through a minimum which balances between accuracy and simplicity. The user can change the default value of CPM in order to optionally control this trade-off. Different values of CPM ranging from 0.01 to 5 were investigated in this study in order to determine the optimum value of CPM which ensures both exactness and generality. After investigating a range of values of CPM the best performing values were considered for building the models. The effects of CPM values on the traffic flows of approach 6 are shown in Figure 4-3. It seems that the CPM values 0.2 and 0.3 make balances between sufficient training and testing error. After obtaining the optimum CPM, the effects of the size of layer 1 are also investigated which is reported in the Figure 4-4. Typically, larger size of layer 1 ensures proper learning of the weights for the training datasets and smaller size of layer 1 will force the model to restrict itself among the fixed inputs which may not have enough predicting capabilities. Figure 4-4 reveals that training and testing error remain constant for 3 to 8 numbers of inputs for the layer 1. The error is also consistently constant for 9 to 12 number of inputs for layer 1. However, the effect can be different depending on the considered CPM value.

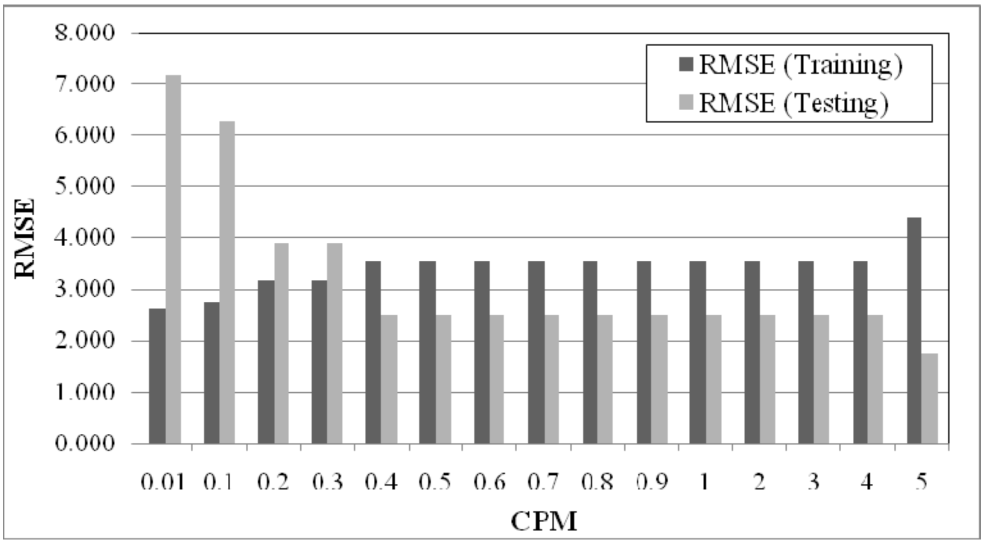


Figure 4-3: The effects of CPM on the prediction of traffic flow of the Approach 6 in the intersection A.

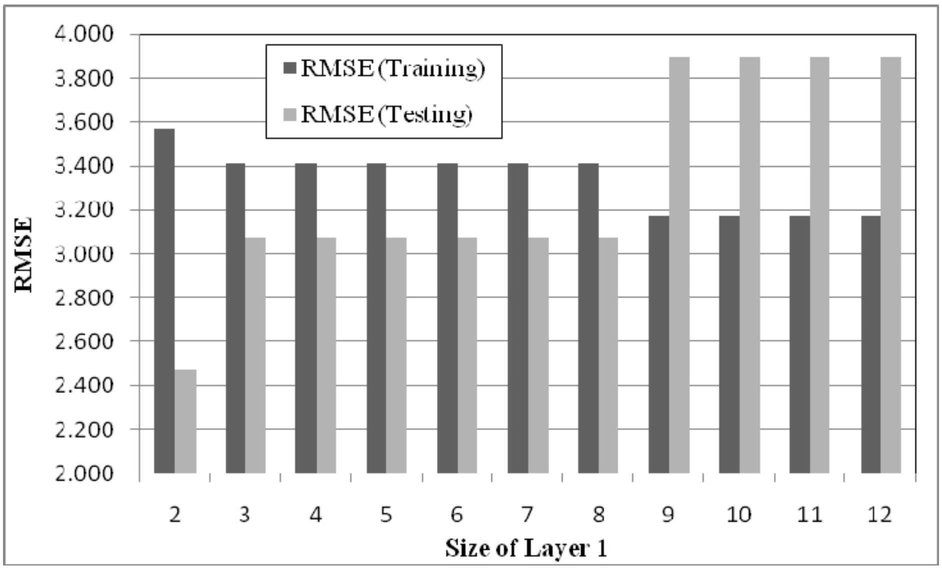
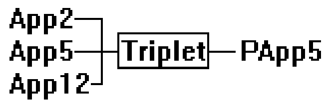
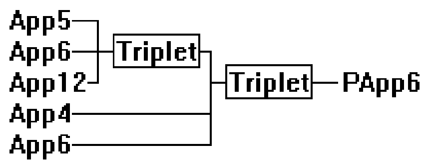
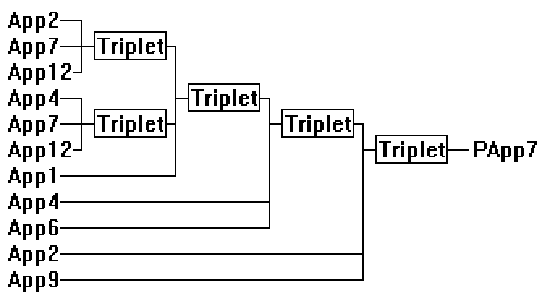
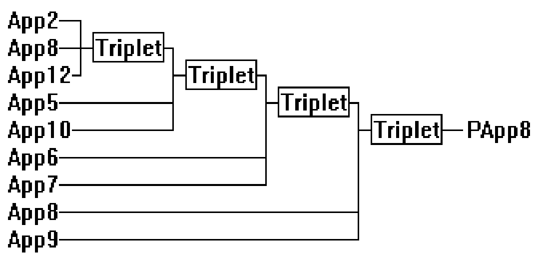


Figure 4-4: The effects of the size of layer 1 for the prediction of traffic flow of the Approach 6 with the CPM value of 0.3.

In this study, four models were developed to predict the 24-hour traffic flows for an average day of a given month expressed in 15-min intervals at intersection “A” using the approach traffic flows of the intersections “A”, “C”, and “B” during the previous month. Therefore, the outputs of the models include the traffic flows of the approaches of intersection A. During the attempt to build the traffic prediction model, the sequential input data were considered. Four networks were developed based on 6 numbers of limiting layers (Table 4-1). In Table 4-1, the “App6” indicates approach number 6. For other approach traffic flows, a similar notation is adopted. It shows the abductive networks which are selected ultimately for predicting the traffic flows of the approaches of the middle intersection “A”. The mentioned CPM values in the Table 4-1 indicate the optimum values which are selected after several trials and errors. Typically, the considered CPM values balance between training and testing errors. The table reveals that the abductive network of Approach 5 is the simplest one among all and those of the Approach 7 and 8 are very complex.

In order to investigate the performance of AIM model compared to NN, feed forward neural networks with two hidden layers having six neurons in each layer were designed in order to predict the approach traffic flows of intersection A. The neural network model of an approach was constructed with the same input used in building the corresponding abductive network model. The activation functions of the neurons of the hidden layers were “tansigmoid” and the activation fuction for the output layer was “purelin”.

Table 4-1: Abductive networks for the approach traffic flows of the intersection A.

Predicted Approach Number	**Abductive Network**	**CPM**
Approach 5		0.2
Approach 6		0.3
Approach 7		0.01
Approach 8		0.01

4.1.6 Performance Evaluation

The predicted values obtained by using abductive models are shown in Figure 4-5, Figure 4-6, Figure 4-7, and Figure 4-8. Figure 4-5 indicates that the model performs well during the off-peak hours. The model output doesn't match with the actual traffic flow during the peak hours especially between 13:00 hour to 17:00 hour. The performance of AIM model for Approach 6 is almost identical to that of Approach 5 (Figure 4-6). During the peak hours between 13:00 hour to 17:00 hour the model is not performing well compared to the other parts of the day. As both the approaching traffic is probably coming from the same type of land-use, their traffic flow characteristics may have similar features which forcing the prediction models to perform accordingly. The AIM model performs consistently well in predicting the traffic flow of Approach 7 (Figure 4-7). But during one peak hour the ANN model performs better than the corresponding AIM model. The AIM model performs consistently well in predicting the traffic flow of Approach 8 (Figure 4-8). It is observed that during one of the peak hours the ANN model performs a bit better than the corresponding AIM model.

In order to investigate the performance of the abductive models, Root Mean Square Error (RMSE), Mean Absolute Error (MAE) of the testing datasets, R-Square value between the actual and predicted traffic flows is selected. The predicted value of the model is matched with the actual value, and the mean differences (D) between them and standard deviation (s) of the differences are also reported.

The error measures are reported in Table 4-2. The abductive models were performing better than the corresponding neural network models except for the approach 6. The R-sq values of the abductive models were varying between 0.855 and 0.991, and those values of the ANN models were varying between 0.830 and 0.988. In terms of RMSE, MAE, and R-Sq values abductive models were performing better than ANN models. The mean of error (D) values of the ANN models is closer to zero compared to the corresponding abductive models. The standard deviation of error (s) values of the ANN models is lower for two approaches which indicate narrower widths of confidence intervals.

The RMSE and MAE of the AIM and NN models are shown in a bar chart (Figure 4-9). The investigations of the charts reveal that RMSE and MAE of Approach 5 and 6 are smaller than that of Approach 7 and 8 for both AIM and NN models. The monthly average and peak of daily traffic of Approach 5 and 6 are very low compared to that of the remaining approaches. The peak 15-min traffic flow of Approach 5 and 6 are around 40 vehicles/15 min but in case of Approach 7 and 8 the values are around 300 vehicles/15 min. If the error measure is normalized with respect to the peak 15-min traffic flow then the performance of the models of Approach 7 and 8 will be better than the models of the remaining approaches which are shown in Figure 4-9.

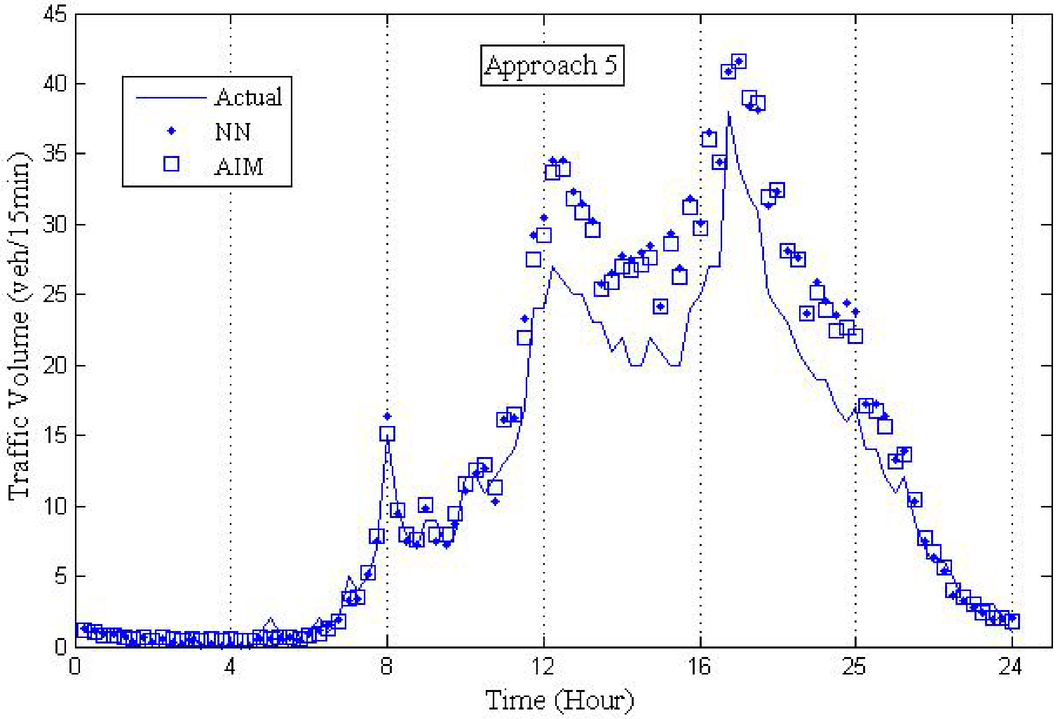


Figure 4-5: The predicted and actual values of the models for the Approach 5.

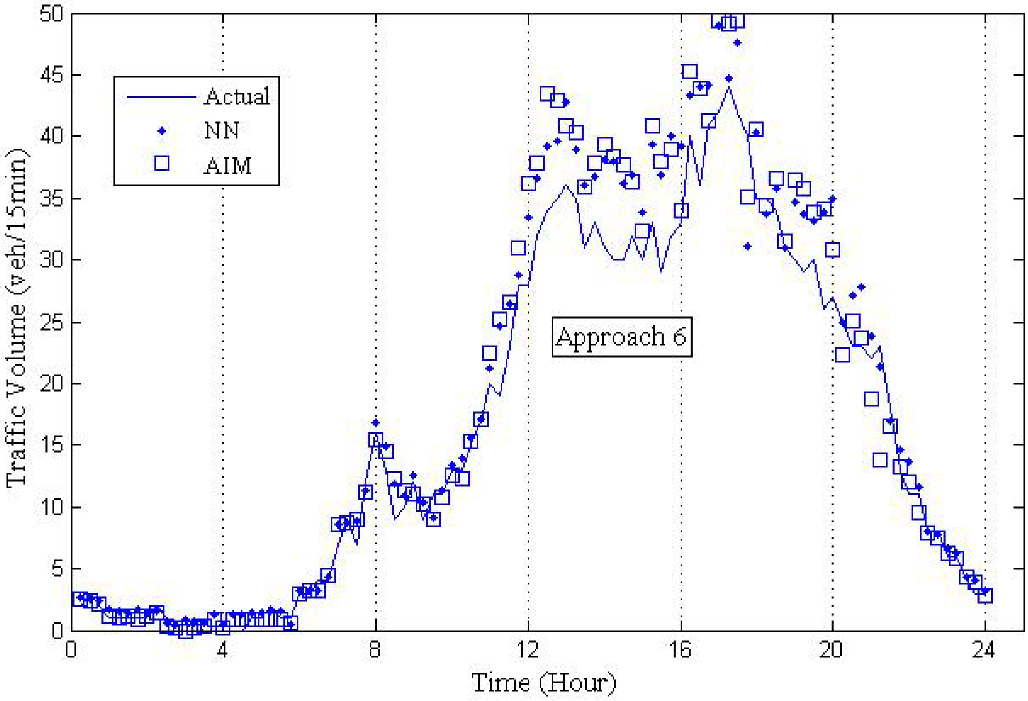


Figure 4-6: The predicted and actual values of the models for the Approach 6.

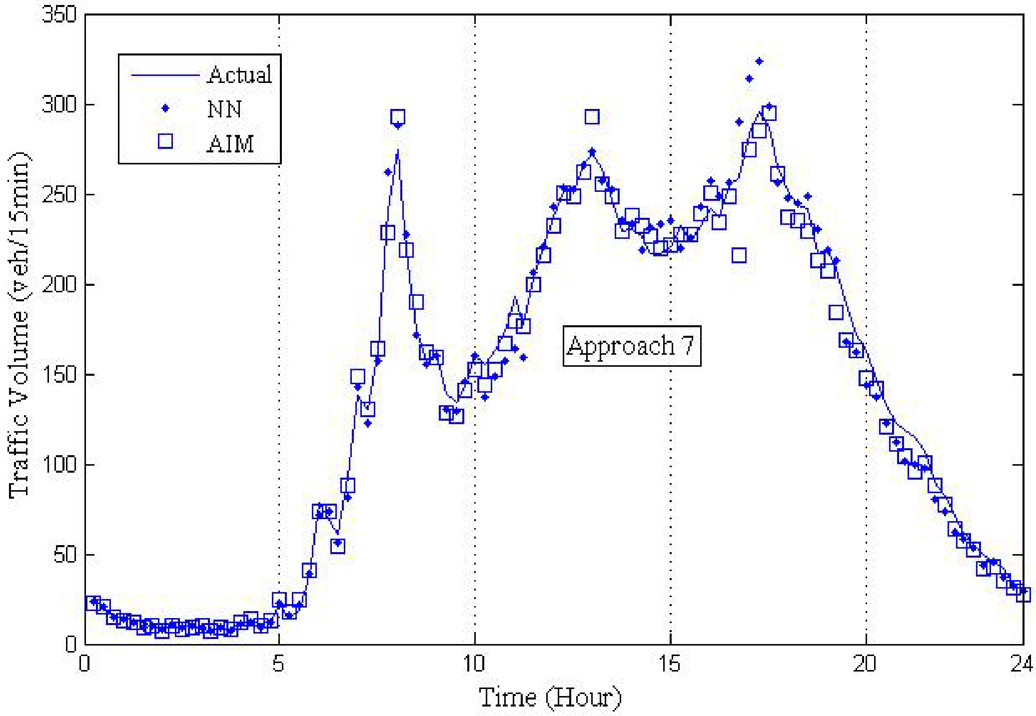


Figure 4-7: The predicted and actual values of the models for the Approach 7.

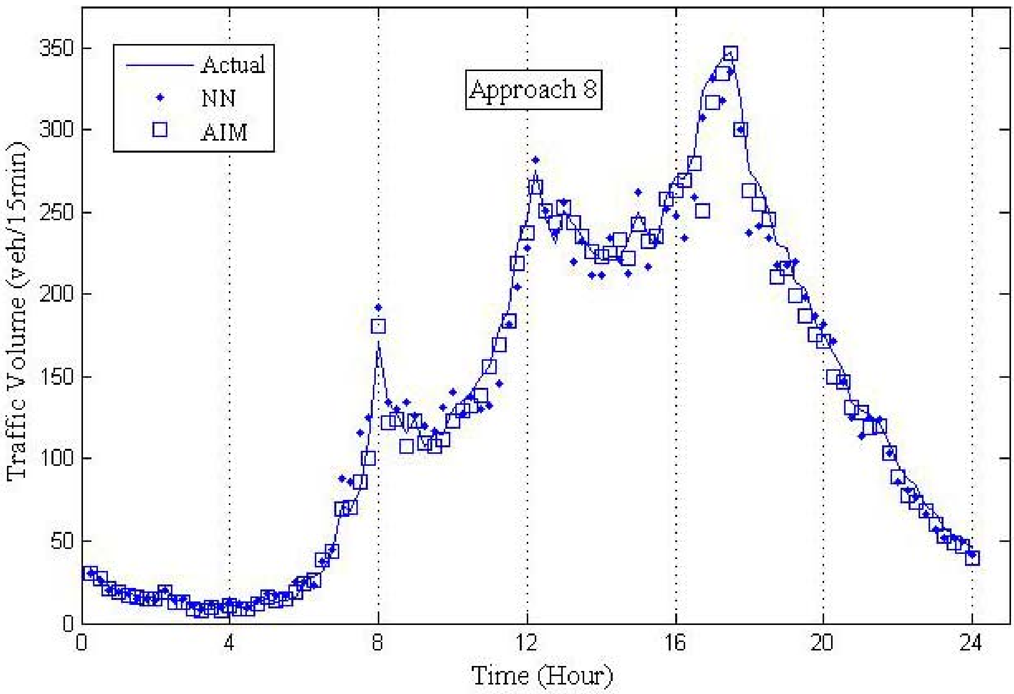


Figure 4-8 : The predicted and actual values of the models for the Approach 8.

Table 4-2: Comparison of prediction accuracy between AIM and NN models of the approaches of the intersection "A".

		RMSE	MAE	R-SQ	Mean (D)	Std. Dev. (s)
Approach 5	NN	4.1762	2.9064	0.8302	-2.4542	2.9951
	AIM	3.8601	2.7156	0.855	-2.5749	3.3052
Approach 6	NN	3.5259	2.4585	0.9351	-2.0571	2.8787
	AIM	3.8952	2.6227	0.9207	-1.8798	3.4296
Approach 7	NN	10.1889	7.1164	0.9884	0.6892	10.219
	AIM	9.0471	6.3732	0.9908	3.1705	8.5178
Approach 8	NN	13.1072	9.3996	0.9829	3.4402	12.7141
	AIM	10.3215	6.0501	0.9894	4.498	9.3386

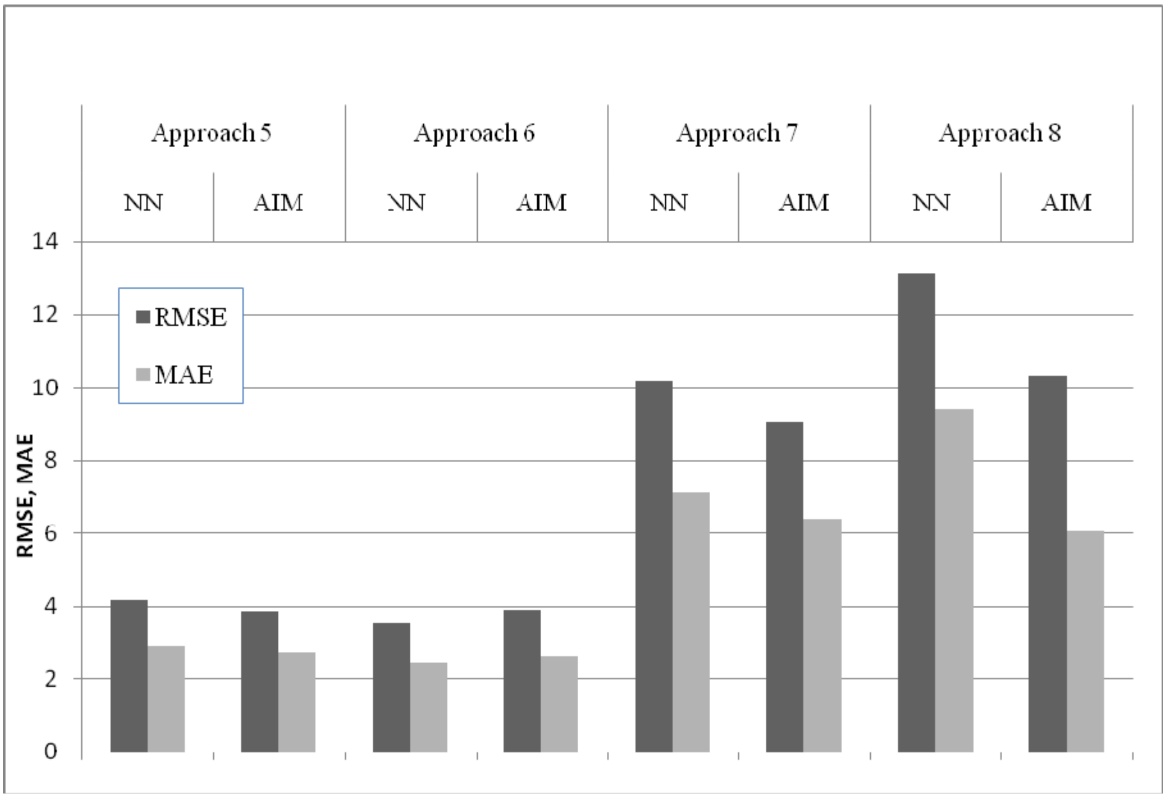


Figure 4-9 : The values of the considered error measures obtained from ANN and AIM models.

4.1.7 Conclusion

This study developed abductive networks of 24-hour traffic flows for an average day of a given month expressed in 15-min intervals. This study exploited the well-proven optimization criteria of AIM in determining network size, element types, connectivity, and coefficients for the optimum model and also investigated the effects of the optionally changeable parameters such as CPM and size of layer 1 on the error measures, aiming to develop best performing networks. The best performing networks were developed based on sequential training datasets with different values of CPM, and 6 numbers of limiting layers. The models also outperformed the ANN models which were built with the same input datasets. The MAE and RMSE are higher for the approach 7 and 8 for both models due to the higher traffic flow compared to other approaches.

Due to the self-organizing nature of the model and minimum required interventions, the proposed AIM model can be easily used by the transportation engineering practitioners to reduce cost and time in traffic volume collection efforts. If it is possible to develop a reliable prediction model for the traffic flows of an intersection, then the traffic flow data can be obtained by using prediction model instead of traffic counter. It will ultimately help in reducing cost and time. It will be useful for rapidly developing countries where comprehensive and continuous traffic data collection is still in its infancy. The availability of the mathematical expression of the model will also enhance the acceptability of the model in larger communities.

4.1.8 Summary of Section 4.1

This section introduces the concept of GMDH which is a representative of inductive procedure which attempts to address the subjective approaches of many AI based models with the help of the principle of self-organization. Then the GMDH based AIM model is illustrated and developed for traffic flow prediction. Although the GMDH algorithm has been successfully used in many identification and forecasting problems, this promising AI technique has not been explored yet in the literature for traffic flow prediction as per the knowledge of the author. The proposed AIM models were evaluated and compared with a feed forward neural network model. The obtained results and the corresponding error measures demonstrate that the AIM model is a valid and promising alternative for predicting intersection traffic flows.

4.2 ADAPTIVE NEURO-FUZZY MODEL

In this study, adaptive neuro-fuzzy models are explored for intersection traffic flow forecasting. Adaptive Neural Network Fuzzy Inference System (ANFIS) is a kind of adaptive neuro-fuzzy model which is considered in this chapter. ANFIS model has been successfully applied in many areas and a few models are also proposed for traffic flow prediction. But most of the ANFIS models are based on subtractive clustering (SC) algorithm which is deterministic in nature. As an alternative, fuzzy C means (FCM) clustering based ANFIS (ANFIS-FCM) model is investigated in this chapter.

This section is divided into six main sub-sections. The next sub-section introduces the fundamentals of ANFIS. The third sub-section describes the adopted procedure in developing the ANFIS models. The fourth sub-section reports the performance of ANFIS-FCM, ANFIS-SC and neural network based forecasting models. Finally, the conclusions and summary are provided in the last two sub-sections.

4.2.1 Concept of Adaptive Neural Network Fuzzy Inference System

Jang (1993) developed ANFIS which serve as a basis for constructing “if-then” rules and Fuzzy Inference System (FIS), and its architecture is obtained by embedding the FIS into the framework of adaptive networks. The ANFIS architecture was successfully used to model nonlinear functions, identify nonlinear components on-linely in a control system, and predict a chaotic time series (Jang, 1993). The neurofuzzy methods provide models which can be interpreted by human beings and the models are in the form of the if-then rules, implying easy integration with operators’ (expert) rules (Janten, 1998). Janten

(1998) stated that the ANFIS model shows good performance, but sometimes produces spurious rules, which makes little sense. This problem can be solved by using an initial FIS generated by clustering techniques such as FCM and SC. FCM clustering algorithm is the soft extension of the traditional C-means which considers each cluster as a fuzzy set, while a membership function measures the degree to which each training vector belongs to a cluster (Tsao et al., 1994). As a result, each training vector may be assigned to multiple clusters which can partly overcome the drawback of dependence on initial partitioning cluster values in hard C-means (Zhang and Chen, 2002).

The structure of ANFIS consists of five-layered feedforward network and the nodes in each layer have the same functionality (Fakhreddine and de Silva, 2004). A typical example of a Sugeno-type ANFIS model is given in Figure 4-10. It consists of five layers. First layer is intended to fuzzify the crisp input by mapping into membership functions. Each node of this layer stores the parameters to define the a bell-shaped membership function. The output of this layer is the value of membership function, $\mu_{A_i^j}(x)$, where x is the input, and A_i^j is the corresponding linguistic label. Each node of the second layer contain each if-then fuzzy rule which performs connective operation "AND" by T-norm operator within the rule antecedent. The output of a node of this layer, w_p indicates the fulfillment of the rule. In the third layer, the relative degree of fulfillment for every rule, $\overline{w}_j$ is obtained with the help of normalization. The fourth layer is concerned with the consequent part of the fuzzy rule and the output of a node of this layer is $\overline{w}_j(\alpha + \beta x_1+\gamma x_2)$, where α, β, and γ are constants. Finally, an output node of the ANFIS model derives the final output by summing all the incoming outputs from the fourth layer. The hybrid learning algorithm of the ANFIS is a combination of the least-squares method and the gradient descent technique for tuning the architecture (Jang, 1993). The premise and consequent parameters are learnt with the help of gradient descent and least-squares method respectively.

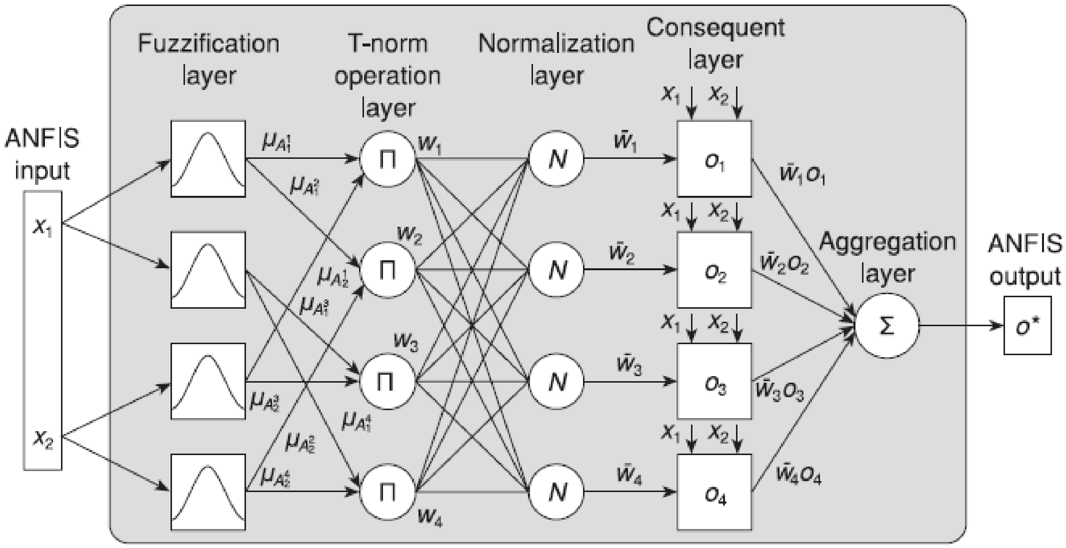


Figure 4-10: An example of a five-layered Sugeno-Type ANFIS model. (source: Fakhreddine and de Silva, 2004)

4.2.2 Model Development

In this subsection, the selection of parameters and network development process for the proposed models are going to be discussed.

In this study, the spatial and temporal data are considered as inputs to forecast the 24-hour traffic flow for an average day of a given month expressed in 15-min intervals. The monthly mean daily 15-min traffic flow of the three intersections are used to forecast the following monthly mean daily 15-min traffic flow of the middle intersection (A). In order to maintain the transparency of the underlying models and reduce the complexity of computation needed for building model, selection procedure is performed to determine the important inputs among all the available candidate inputs. The purpose of input selection is to remove noise/irrelevant inputs, remove inter dependent inputs, make the underlying model more concise and transparent, and reduce the time for model construction (Jang, 1996). In ANFIS, the least-squares method leads to fast training and the gradient descent method slowly changes the underlying membership function that generates the basis functions for the least-squares methods (Jang, 1996). Therefore, it can be expected that ANFIS may produce satisfactory results after a few epochs of training. The inputs are selected by adopting the approach of Jang (1996). ANFIS models were built by gradually increasing the number of inputs with respect to the RMSE (root mean squared error) obtained after few epochs of training. Table 4-3 shows the selected inputs and number of membership functions (MFs) for the forecasting model of the approach traffic flows of the intersection A. The selected inputs ensure better performance and any addition or subtraction of input increases the values of error measurements. Typically, increased numbers of MFs and rules ensure better training but they reduce the generalization capability of the model.

Table 4-3: Selected inputs and number of MFs of each input for ANFIS models.

Forecasting Model	Input	Number of MFs of each input	
		ANFIS (SUB)	ANFIS (FCM)
Approach 5	Approach 4,5,7,9 and 10	4	3
Approach 6	Approach 3,4,5,6,7	3	2
Approach 7	Approach 3,4,5,7	3	3
Approach 8	Approach 7,8,9	4	2

The direct implementation of ANFIS causes spurious number of rule generation. In order to solve this problem, the model was initialized with a number of FIS obtained by using FCM clustering and subtractive clustering methods. In this study, only generalized bell shaped MFs for inputs were tried because Nayak et al. (2004) concluded that the choice of MFs of ANFIS architecture doesn't have significant impact on the model performance.

In order to build forecasting model, the total datasets were divided into two datasets which include the training and the testing datasets. The training datasets consider the monthly mean 24-hours 15-min traffic flow data starting from July, 2007 to April, 2009. The testing datasets include the monthly mean 24-hours 15-min traffic flow data starting from May, 2009 to June, 2009. Each month traffic flow data of the three intersections (A, B and C) were used to forecast the next month traffic flow data of the intersection A. For the ANFIS models the input data were normalized in such a way that the input varies between 0 and 1.

A feedforward neural network model was also built to evaluate the comparative performance of the ANFIS models. The data were preprocessed to be mapped with mean and standard deviation of 0 and 1 respectively. The model consists of two hidden layers having 6 neurons in each layer. The network was trained for 100 epochs for all the models of the considered approaches. The final topology of the network is achieved after considering preprocessing, number of neurons, number of hidden layers, activation function, training algorithm, etc.

4.2.3 Performance Evaluation

In order to compare the forecasting accuracy of the ANFIS models and the ANN, RMSE and R-square value are reported in Table 4-4. The table shows that ANFIS-FCM model has smaller RMSE values than ANFIS-SC and ANN for all approaches except approach 6. ANFIS-FCM also outperforms other models in terms of R-square values. It simply means that the performance of ANFIS-FCM model is better than that of other models. The predicted values of the models is matched with the actual value and the mean difference (D) and standard deviation of the difference are also reported in

Table 4-4. The mean of differences show that the values of D of ANFIS-FCM model are close to zero compared to other models. The standard deviation (s) of D of ANFIS-FCM is smaller compared to other models which indicate narrower widths of confidence interval.

The RMSE and MAE of the models are shown in a bar chart (Figure 4-11). The investigations of the chart reveals that RMSE and MAE of Approach 5 and 6 are smaller than that of Approach 7 and 8. The monthly average and peak of daily approach traffic of Approach 5 and 6 are significantly smaller than that of the remaining approaches. For example, the peak 15-min traffic flow of Approach 5 and 6 are around 40 vehicles/15 min but in case of Approach 7 and 8 the values are around 300 vehicles/15 min. If the error measures are normalized with respect to the peak 15-min traffic flow then the performance of the models of Approach 5 and 6 will be worse than the models of the remaining approaches. If the error measures are normalized with respect to the total traffic flow along each approach, then the result will be the same. It means the models of Approach 7 and 8 perform better than the models of the remaining approaches.

Table 4-4: Comparison of forecasting accuracy for different models of the approaches of intersection A.

		RMSE	MAE	R-SQ	MEAN (D)	STD DEV (s)
Approach 5	ANFIS-FCM	2.5209	1.6842	0.9381	-1.3465	2.1424
	ANFIS-SC	2.8093	1.9445	0.9232	-1.3893	2.4545
	NN	2.8751	2.0531	0.9195	-1.4344	2.5048
Approach 6	ANFIS-FCM	2.0646	1.3810	0.9777	-0.5158	2.0096
	ANFIS-SC	3.5028	2.1561	0.9359	-1.0816	3.3491
	NN	4.3975	2.2862	0.8990	-1.3229	4.2158
Approach 7	ANFIS-FCM	10.5642	8.2039	0.9875	7.6994	7.2713
	ANFIS-SC	11.4579	8.5659	0.9853	7.9801	8.2652
	NN	10.6439	8.3531	0.9873	7.8968	7.1742
Approach 8	ANFIS-FCM	10.1924	7.5266	0.9897	3.0118	9.7883
	ANFIS-SC	10.3194	7.2747	0.9894	2.4053	10.0879
	NN	13.7558	8.3066	0.9812	3.8920	13.2630

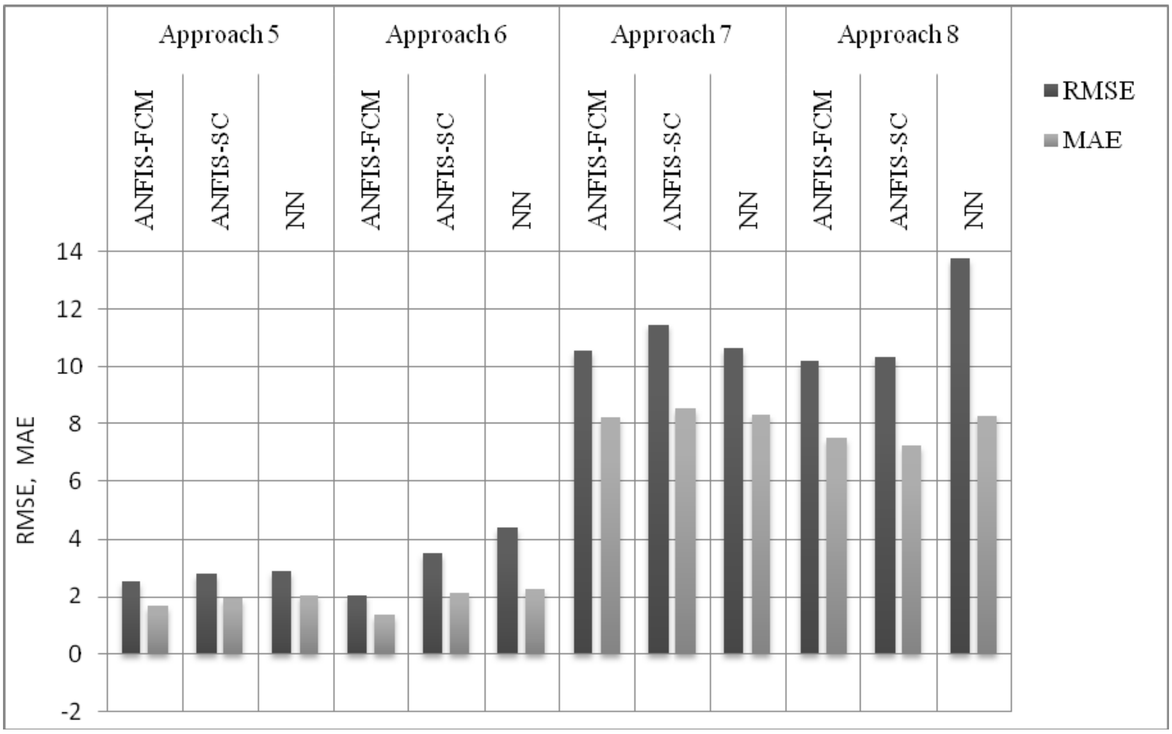


Figure 4-11: Comparison of forecasting accuracy for different models of the approaches of intersection A.

Figure 4-12, Figure 4-13, Figure 4-14, and Figure 4-15 show the predicted output of ANFIS-FCM, ANFIS-SC and NN along with the corresponding actual values. The visual analysis of the output of the models of Approach 5 indicates that ANFIS-FCM performs better than other models. But all the models fail to predict the period between 13:00 hour and 16:00 hour (Figure 4-12). The ANFIS-FCM model of Approach 6 learns the trend almost perfectly for all over the period (Figure 4-13). Through visual inspection it seems that the performance of ANFIS-FCM and ANFIS-SC model for Approach 7 is comparable and both the models can perfectly predict the two peaks out of three (Figure 4-14). Both ANFIS-FCM and ANFIS-SC models perform well in predicting the traffic flow of Approach 8 (Figure 4-15). The models successfully learn the general trend for both low and high traffic flow conditions. The output of ANFIS-FCM and ANFIS-SC for the peaks is very close to the actual values for two cases. In one case, the ANN performs better than both ANFIS models.

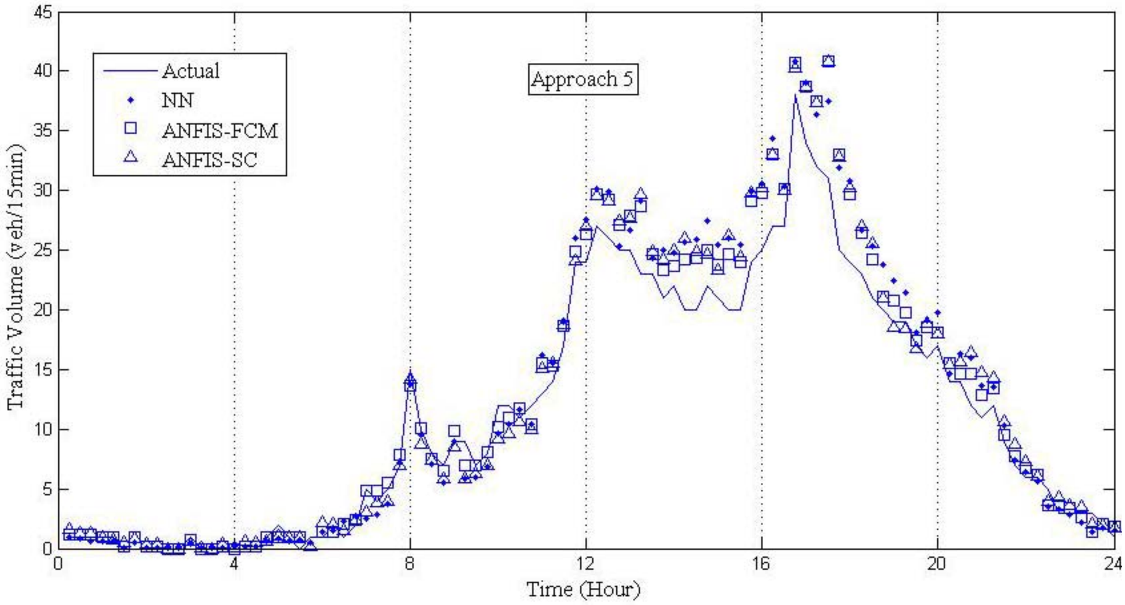


Figure 4-12: The result of the models and the actual value for the approach 5.

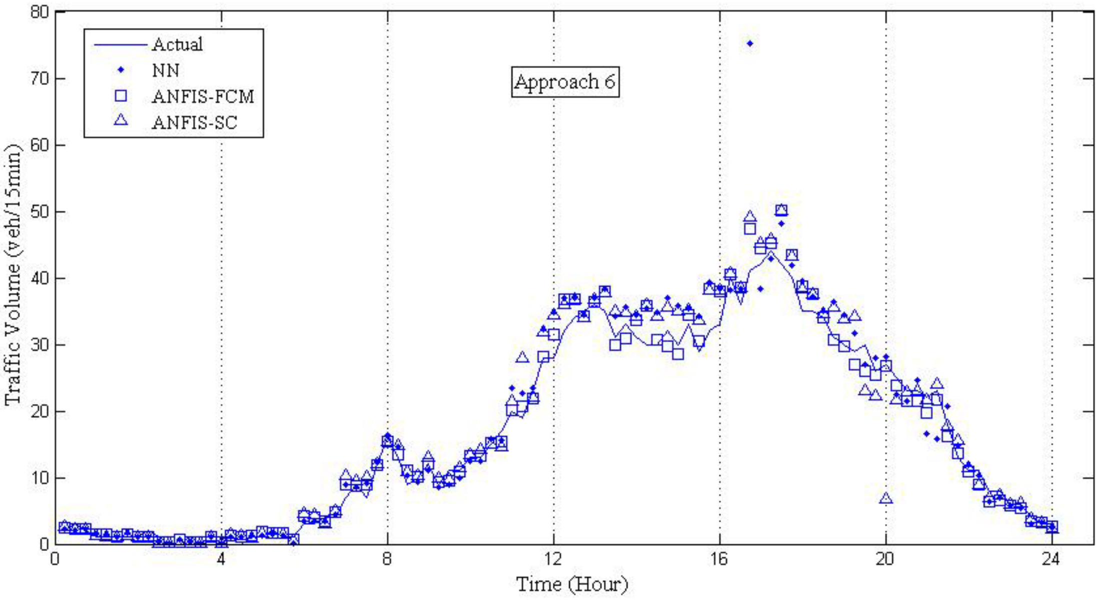


Figure 4-13: The result of the models and the actual value for the approach 6.

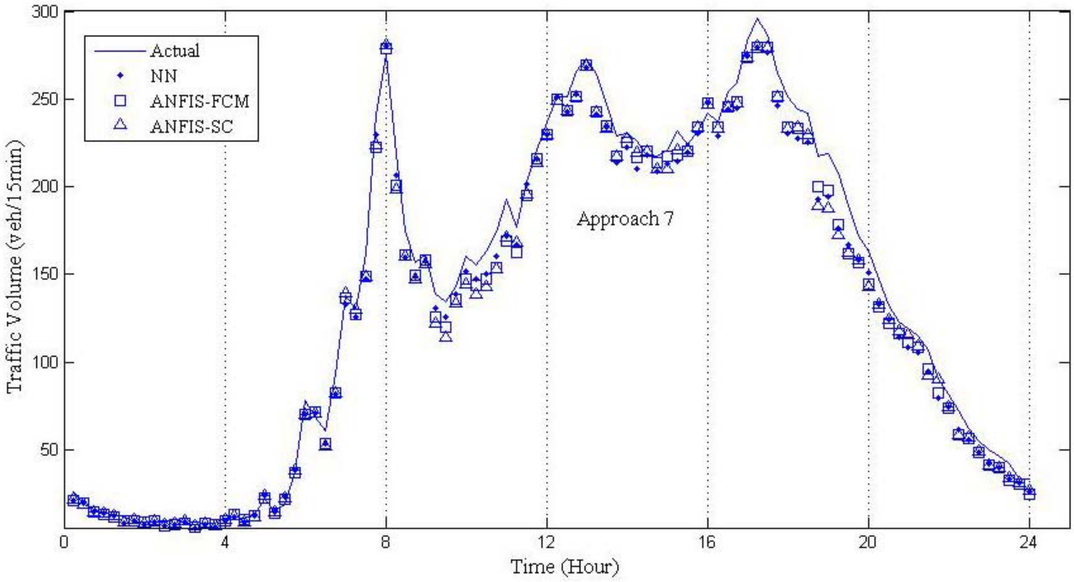


Figure 4-14: The result of the models and the actual value for the approach 7.

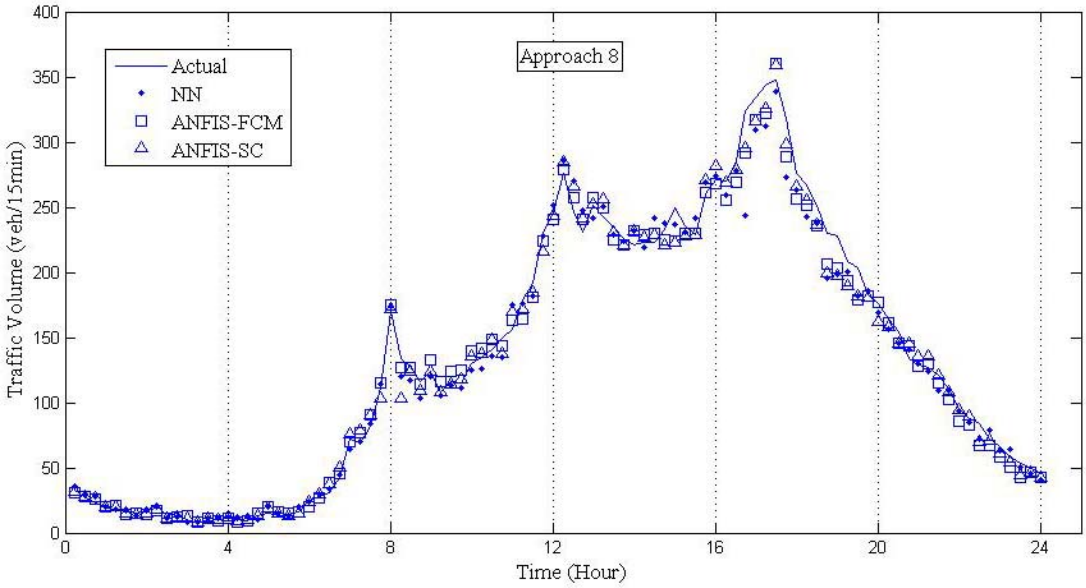


Figure 4-15: The result of the models and the actual value for the approach 8.

4.2.4 Conclusion

Accurate and reliable forecast of intersection traffic flow can help reducing cost of traffic studies and increasing accuracy of traffic studies which include highway geometry, estimates of road revenue, selection of highway routes, selecting the timing of maintenance, signal timing, location and design of highway systems, air quality analysis, design of traffic control systems and accident rates, average daily traffic, location of service areas etc. This study adopted FCM based ANFIS model to predict intersection traffic flow and compared the performance measures of ANFIS-FCM with ANFIS-SC and ANN models. The MAE and RMSE are higher for the approach 7 and 8 due to the higher traffic flow compared to other approaches. The obtained empirical results by ANFIS-FCM yield more accurate intersection traffic flow forecasting than the other models and demonstrate that ANFIS-FCM is a valid and promising alternative for predicting intersection traffic flows. The output of this model is intended to be used for determining TOD breakpoints required for pre-timed or actuated traffic signal controllers.

4.2.5 Summary of Section 4.2

Traffic flow prediction of the intersection play important role in planning, design, and operations of transportation facilities. It is difficult to model the complicated, non-linear relationships of the traffic flow using conventional approaches. This study considered spatial and temporal characteristics, and developed ANFIS based intersection traffic flow model which is initialized with fuzzy inference system obtained by using FCM and SC methods. FCM based ANFIS model is not explored in the literature in the area of intersection traffic flow prediction. The proposed FCM based ANFIS model outperforms a SC based ANFIS and feed forward neural network models. The forecasted 24-hour traffic flows for an average day of a month divided into 15-min intervals are intended to be used for determining TOD breakpoints required for pre-timed or actuated traffic signal controllers.

4.3 TYPE-2 FUZZY LOGIC MODEL

In this section, spatial and temporal characteristics of traffic flows are considered to develop type-2 fuzzy logic based prediction model. The type-2 fuzzy logic system (FLS) has more design degrees of freedom associated with a type-2 FLS compared to a type-1 FLS which indicates higher potential of better performance. The adoption of singular value decomposition (SVD) ensures only important rules in the rule base and saves training time to build the model without any significant loss of accuracy. Type-2 fuzzy models are not explored in the literature in the area of intersection traffic flow forecasts.

The current research is still limited to type-2 fuzzy sets because the complexity of fuzzy logic system increases rapidly with increasing types. The fundamentals of FLS are same for type-1 to type-n fuzzy sets (Lee and Lee, 2002). Karnik and Mendel (1998b) solved the major obstacles to extend a type-1 FLS to a type-2 fuzzy FLS by characterizing type-2 fuzzy sets, performing operations with type-2 fuzzy sets, inferring with type-2 fuzzy sets and provided a defuzzified output from a type-2 inference engine.

4.3.1 Fuzzy Model Building

Fuzzy modeling linguistically specifies approximate relationships between the input and the output which is a very useful approach when the system is complex and poorly understood. Fuzzy models can be built by encapsulating the expert knowledge as fuzzy "if-then" rules and producing rules with the help of heuristic analysis of a system (Zadeh, 1973; Takagi and Sugeno, 1985; Yager and Filev, 1994). Due the complexity and ill-definition of the system, there are significant difficulties in heuristic approach for rule construction. The researchers focus on automatically rule generation techniques. The automatic rule generation approaches can be materialized with the help of neuro-fuzzy systems, evolutionary rule generation, clustering, and proximity analysis (Sudkamp, Knapp, and Knapp, 2003).

Fuzzy model aims to ensure interpretability and precision which are conflicting in nature. The former in this case the interpretability can be obtained through small number of rules having higher generalization capability while the latter in this case the precision can be generally achieved by increased number of rules. Both of these objectives can be achieved by constructing a rule base using a learning algorithm and then by removing the number of rules which are not dominant. The number of rules can be reduced with the help of dimension reduction and rule merging. Singular value decomposition is one of the most popular approaches in the literature for dimension reduction which identifies input variables which may be eliminated (Sudkamp et al., 2003).

4.3.2 Singleton and Non-Singleton Type-1 FLS

The fuzzy models can be built based on singleton and non-singleton Type-1 fuzzy logic systems. Singleton Type-1 FLS is based on the assumptions that all the fuzzy sets are type-1 as there are no uncertainties and measurements are perfect and these are treated as crisp values. The singleton fuzzifier maps a crisp point into a fuzzy set and a fuzzy singleton, A_x has a support x' if membership grade, $\mu_{A_x}(\boldsymbol{x}) = \mathbf{1}$ for x = x' and $\mu_{A_x}(\boldsymbol{x})$ =0 for all other $x \in \boldsymbol{X}$ with $x \neq x'$. The fuzzification of singleton inference system is very easy to evaluate due to the simplification of the sup-star composition. If the data are corrupted by measurement noise then singleton fuzzification may not be adequate.

In a non-singleton type-1 FLS, inputs are modeled as type-1 fuzzy number which can handle the uncertainty in the inputs due to noisy measurements (Mendel, 2001). Both singleton and non-singleton type-1 FLS can be described by Figure 4-16. A non-singleton fuzzifier maps a crisp measurement x = x' into a fuzzy set (Kaufman and Gupta, 1991). In this fuzzification, $\mu_{A_x}(\boldsymbol{x}) = 1$ for x = x' and $\mu_{A_x}(\boldsymbol{x})$ decreases from unity as x moves away from x'. It implies that the existing measurement is the most likely value to be the correct one from all the remaining values. The presence of noise with the measurement causes to consider the neighboring points as the likely correct value with a lesser degree (Mendel, 2001). Depending on the kind and quantity of the existing noise associated with the measurement, the designer can determine the membership function of the measurement.

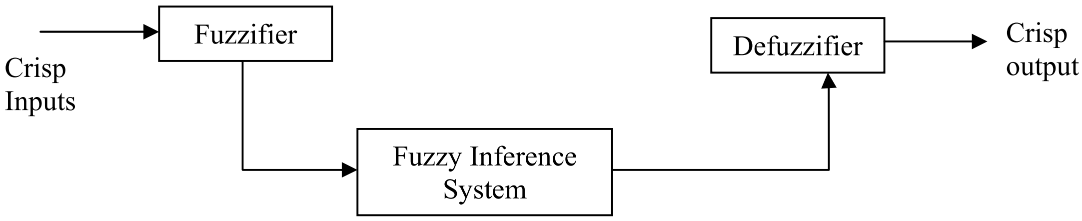


Figure 4-16: Singleton or non-singleton type-1 fuzzy logic system.

4.3.3 Characteristics of Type-2 Fuzzy Sets

In order to explain type-2 fuzzy sets it is required to start with type-1 fuzzy sets. A type-1 fuzzy set, A is characterized by a membership function, $\mu_A(y)$; i.e. $A = \int_{y \in Y} \mu_A(y)/y$, where the integral sign indicates logic union and it is replaced by a summation sign if Y is discrete. A type-2 fuzzy set can be characterized by a three dimensional fuzzy set and the concept of footprint of uncertainty (Karnik et al., 1999), along with the associated concepts of lower and upper membership functions (Mendel and Liang, 1999). A type-2 fuzzy set in Y is denoted by $\tilde{A}$, is characterized by a type-2 membership function $\mu_{\tilde{A}}(y, u)$, where $y \in Y$ and $u \in J_y$, which is a type-1 fuzzy set in [0,1]. The amplitude distribution of all $y \in Y$ can create a three dimensional membership function which is basically a type-2 membership function. The footprint of uncertainty (FOU) of $\tilde{A}$ is a bounded region and is the union of all primary membership grades, i.e., $FOU(\tilde{A}) = \cup_{y \in Y} J_y$. The concept of FOU is very useful because it focuses the attention on inherent uncertainties in a type-2 membership function and it provides a convenient verbal description of the entire domain of support for all the secondary grades of a type-2 membership function (Mendel, 2001). The upper and lower membership functions of $\tilde{A}$ are associated with the upper and lower bound of $FOU(\tilde{A})$ respectively.

4.3.4 Operations Used in Type-2 Fuzzy Sets

Although the common operations of union, intersection, and complement for type-1 fuzzy set are well-developed for quite long time, these operations for type-2 fuzzy sets were developed mainly by the contributions of Karnik and Mendel (2001a, 2001b). The union of two type-2 fuzzy sets, $\tilde{A} = \int_u f_y(u)/u$ and $\tilde{B} = \int_v f_y(v)/v$ is another type-2 fuzzy set whose MF can be computed by adopting the following equation.

$$\mu_{\tilde{A} \cup \tilde{B}}(y) = \int_u \int_v f_y(u) * g_y(v)/(u \vee v)$$

where, $\cup$ is the join operation and * is the selected t-norm (product or minimum). The meet operation ($\cap$) of two type-2 fuzzy sets is shown in the following equation. The detail of join and meet operation is provided in Karnik and Mendel (1998a; 1998b).

$$\mu_{\tilde{A} \cap \tilde{B}}(y) = \int_u \int_v f_y(u) * g_y(v)/(u \wedge v)$$

The complement of a type-2 fuzzy set, $\tilde{A}$ is an another type-2 fuzzy set which can be obtained as follows.

$$\mu_{\overline{\tilde{A}}}(u) = \int_u f_y(u)/(1-u)$$

A fuzzy knowledge base represent a complex system as a simple set of input-output rules and the design of the rule base is an important step of it. In fuzzy knowledge base systems, the rules are fuzzy relations and the collective set of fuzzy relations form the knowledge base. The fuzzy relation formed by the collection of rules can be presented by a multivariable membership function. The degree of presence or absence of association, interaction, or interconnectedness between the elements of two or more fuzzy sets can be represented by fuzzy relations (Mendel, 2001) and play an important role in a fuzzy logic system. A type-2 fuzzy relation $\tilde{F}(A_1, \dots\dots, A_m)$ is a type-2 fuzzy set which is defined on the Cartesian product space of crisp sets $A_1, \dots\dots, A_m$.

$$\tilde{F}(A_1, \dots\dots, A_m) = \int_{A_1 X A_2 X \dots X A_m} \mu_{\tilde{F}}(a_1, a_2, \dots\dots, a_m)/(a_1, a_2 \dots\dots, a_m) \quad \text{and } a_i \in A_i$$

In a fuzzy decision making process, the knowledge-base is collectively matched with the available data and then an inference is made on another fuzzy variable based on the result of the matching process. Composition operation performs this matching and inference-making and plays a crucial role in fuzzy inference and fuzzy knowledge-based systems. A type-2 FLS having a rule base of M rules in which each rule has q antecedents, let the *n*th rule be represented by R^n, where R^n: IF y_1 is $\tilde{F}_1^n$, y_2 is $\tilde{F}_2^n$, ………., and y_q is $\tilde{F}_q^n$, THEN z is $\tilde{G}^l$. The following extended sup-star composition expresses the membership function: $\mu_{\tilde{D}^n}(z) = \bigcup_{y \in Y}\left[\mu_{\tilde{C}_y}(y) \cap \mu_{\tilde{C}^n \to \tilde{D}^n}(y,z)\right]$ where Y is q-dimensional Cartesian product space, $Y = Y_1 \times Y_2 \ldots \times Y_q$, Y_k is the measurement domain of input y_k, $(k = 1, \ldots, q)$; and $\tilde{A}_y$ is represented by $\mu_{\tilde{A}_y}(y) = \mu_{\tilde{Y} \times \ldots \times Y^q} = \mu_{\tilde{Y}_1}(y_1) \cap \ldots \cap \mu_{\tilde{Y}_q}(y_p)$.

Additionally, $\mu_{\tilde{C}^n \to \tilde{D}^n}(y,z) = \mu_{\tilde{F}_1^n}(y_1) \cap \mu_{\tilde{F}_2^n}(y_2) \cap \ldots \cap \mu_{\tilde{F}_q^n}(y_p) \cap \mu_{\tilde{G}^n}(z)$

4.3.5 Applications of Type-2 Fuzzy Logic Model

Type-2 FLS has been successfully applied in wide range of areas including classification of video streams (Liang and Mendel, 2000a), co-channel interference elimination from communication channels (Liang and Mendel, 2000b), connection admission control (Liang et al., 2000), control of mobile robots (Wu, 1996), decision making (Chaneau et al., 1997; Yager, 1980), solving fuzzy relation equations (Wagenknecht and Hartmann, 1988), survey processing (Karnik and Mendel, 1999b), time-series forecasting (Karnik and Mendel. 1999a), and preprocessing of data (John et al., 1998).

4.3.6 Methods of Rule Reduction: Singular Value Decomposition

Fuzzy logic system modeling suffers from the problem of the curse of dimensionality because the number of rules increases exponentially with the increase of the number of input variables (Yen and Wang, 1996). In order to solve the problem of rule explosion in both type-1 and type-2 FLS, many design approaches have been proposed in literature. Wang et al. (1995) developed a fuzzy model reduction approach based on the optimality theorem of Johansen (1995). In this approach, a subset of fuzzy rules are selected which satisfy the condition of completeness (Willaeys and Malvache, 1979). This approach may miss some dominant fuzzy rules and sometimes lead to a large loss of accuracy (Yen and Wang, 1996). Takagi and Sugeno (1985), and Chiu (1995) focused on selecting dominant input variables by adopting heuristic-type algorithm. Yen et al. (1993) proposed principal component analysis to reduce input dimension. Sugeno et al. (1993) developed a hierarchical-structured fuzzy controller which ensures linear increase of the number of fuzzy rules with the increase of the number of input.

The current trend seems to focus on orthogonal transformation methods for selecting important rules from a given rule base. Yen and Wang (1996) utilized SVD to detect and select the dominant fuzzy rules from a rule base and illustrated the effectiveness by using a nonlinear limit cycle modeling and prediction problem.

SVD is developed based on the idea that the singular values can be used to decompose a given system and indicate the degree of significance of the decomposed parts (Baranyi et al., 2003). Reduction is obtained by truncating the parts which have weak or no contribution at all to the output depending on the assigned singular values (Baranyi et al., 2003). SVD is a very powerful and useful tool for modern numerical analysis and has been applied successfully in statistical analysis (Hammarling, 1985), image processing (Andrews and Patternson, 1976), signal processing (Comon and Golub, 1990; van der

Veen et al., 1993), control (Laub, 1985), and system identification (Vandewalle and De Moor, 1988).

Yen and Wang (1999) attempted to solve the same kind of problem by investigating various techniques such as orthogonal lease-squares, eigen value decomposition, singular value decomposition (SVD-QR) with column pivoting method, total least-squares method, and direct SVD method. Mouzouris and Mendel (1996, 1997) proposed SVD-QR method for determining important fuzzy rules. SVD is used as both fuzzy rule base reduction and structure decomposition (Yam et al., 1999; Baranyi and Yam, 1997). Depending on the SVD of the rule consequents, Yam et al. (1999) generated certain linear combinations of the original membership functions for a reduced set of fuzzy rules.

Liang and Mendel (2000c) proposed a design method for interval type-2 FLSs but it doesn't solve the problem of rule explosion. Liang and Mendel (2000d) proposed a SVD-QR method based design procedure to address the problem of rule explosion.

The singular value decomposition of an N-by-M matrix B is a factorization of B into a product of three matrices which is as follows.

$B = V \sum U^T$ When $V^T V = U^T U = I_R$ and $\sum = diag(\sigma_1, \sigma_2, \ldots., \sigma_r)$

Where, $R = \min(N, M)$

The elements of the matrices V and U are the orthonormalized eigenvectors associated with the R largest eigenvalues of BB^T and B^TB, respectively and the diagonal elements of $\sum$ are the non-negative square roots of the R largest eigenvalues of BB^T and are known as singular values.

4.3.7 Proposed Type-2 Fuzzy Logic Model

Fuzzy modeling linguistically specifies approximate relationships between the input and the output which is a very useful approach when the system is complex and poorly understood. The antecedent of the fuzzy "if-then" rules model the state of the underlying system and the consequent produces the potential responses. Fuzzy models can be built by encapsulating the expert knowledge as fuzzy rules and producing rules with the help of heuristic analysis of a system (Zadeh, 1973; Takagi and Sugeno, 1985; Yager and Filev, 1994).

There are many methods to automatically construct fuzzy systems by using numerical input-output data. The adopted FLS in this study is shown in Figure 4-17. In this approach all the fuzzy input sets and the random center of consequents are used to build the initial rules which are equal to the sample size. In the next step, the SVD technique is applied to reduce the number of rules and determine the most important rules to produce desired output. Then obtained non-singleton type-1 FL model is used to initialized the parameters of type-2 FL model. The type-2 model is fine-tuned with the help of back-propagation algorithm.

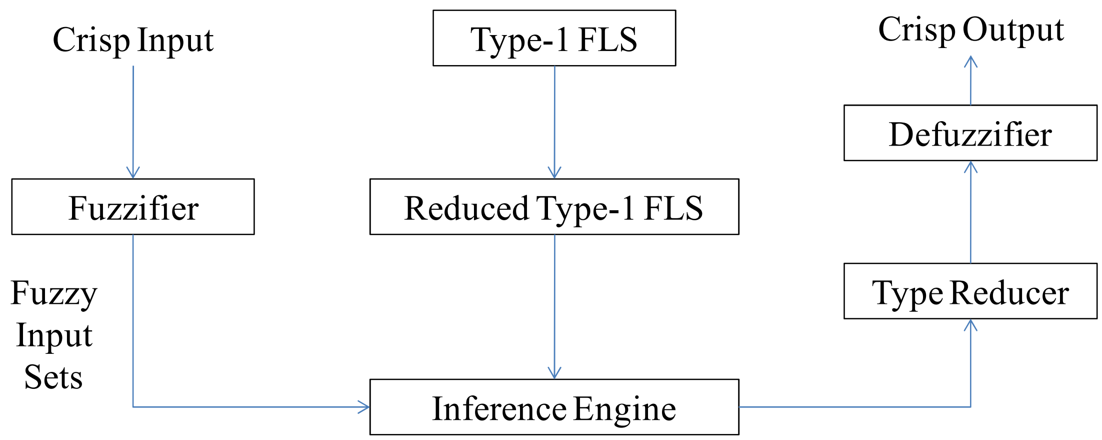


Figure 4-17: SVD based Type-2 Fuzzy Logic System.

In a type-2 FLS, it is required to decide on the kind of fuzzification (either singleton or non-singleton), select a FOU for membership function, decide functional forms for both the primary and secondary membership functions, select the initial parameters of the membership function. It is also needed to predefine the type of composition, implication, type-reduction method and defuzzifier. In this study, all the proposed non-singleton type-2 models assume Gaussian primary membership functions with uncertain mean and interval secondary membership functions, product implication and t-norm, center-of-sets type reduction, and defuzzification obtained by the centroid of the type-reduced set.

Table 4-5 shows the input selected for both type-1 and type-2 FLS. The models of Approach 5 and 6 only require two inputs. The model of Approach 7 requires 5 inputs and that of Approach 8 requires only one input. It seems that FLS require less number of inputs compared to AIM and ANIFS models which are described in the previous sections.

Table 4-5: Selected inputs for type-1 and type-2 fuzzy models.

Prediction Model	Input
Approach 5	Approach 2 and 5
Approach 6	Approach 5 and 6
Approach 7	Approach 3,7,9,10 and 11
Approach 8	Approach 8

4.3.8 Performance Evaluation

Figure 4-18, Figure 4-19, Figure 4-20, and Figure 4-21 show the predicted output of the non-singleton type-2, non-singleton type-1, and singleton type-1 model. It seems that the non-singleton type-2 models can learn the general trend of the intersection traffic flow including the peak hours better than the other models. In almost all the cases, the models provide overestimation compared to the actual values.

The type-2 model predict the traffic flow of Approach 5 accurately without losing the general trend of the actual values except for a short duration between 12:00 and 16:00 hour (Figure 4-18). For very low traffic flow condition at the early morning, non-singleton type-1 performs better than type-2 model.

The performance of type-2 model is better than type-1 models in predicting the traffic flow of Approach 6. But for the very low traffic flow condition at the early morning, both types of type-1 models outperform type-2 model (Figure 4-19). There are three predominant peak 15-min traffic flows for the Approach 6. For two peaks, type-1 models perform better than type-2 model.

The performance of type-2 model is better than other type-1 models for the Approach 7 and 8 for almost all the points (Figure 4-20 and Figure 4-21). Type-2 performs well for both low and high traffic flow conditions.

In order to compare the prediction accuracy of the type-2 fuzzy logic model and the singleton and non-singleton type-1 FLS RMSE, MAE and R-square values are reported in Table 4-6. The table shows that type-2 model has smaller RMSE, MAE values than singleton and non-singleton type-1 models. Type-2 model also outperforms other models in terms of R-square values. It indicates that the performance of type-2 model is better than other considered models in this chapter. Non-singleton type-1 model outperforms the singleton type-1 model with respect to RMSE, MAE and R-square values.

The predicted values of the models are matched with the actual values to determine the differences. The mean difference (D) and standard deviation (s) of the differences are reported in Table 4-6. The table shows that the values of "D" of type-2 models are close to zero compared to other models for all the approaches. The "s" values of type-2 models are the smallest compared to other models which indicate narrower widths of confidence interval.

The RMSE and MAE of the type-1 and type-2 models are shown in a bar chart (Figure 4-22). The investigation of the chart reveals that RMSE and MAE of Approach 5 and 6 are smaller for singleton and non-singleton type-1 models, and non-singleton type-2 model compared to Approach 7 and 8. It is observed that the monthly average and peak of daily approach traffic of Approach 5 and 6 are significantly smaller than that of the remaining approaches. For an instance, the peak 15-min traffic flow of Approach 5 and 6 are around 40 vehicles/15 min but in case of Approach 7 and 8 the values are around 300 vehicles/15 min. If the error measures are normalized with respect to the peak 15-min traffic flow then the performance of the models of Approach 7 and 8 will be better than the models of the remaining approaches.

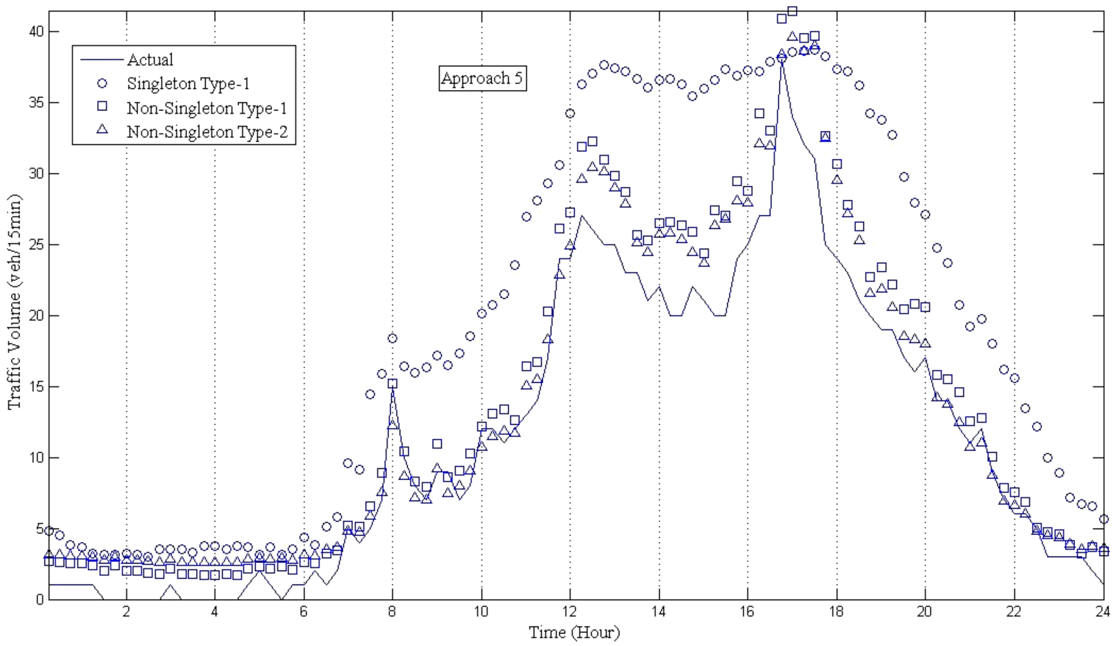


Figure 4-18: The predicted output of singleton and non-singleton type-1 FLS and non-singleton type-2 FLS, and the actual value for the Approach 5 at intersection A.

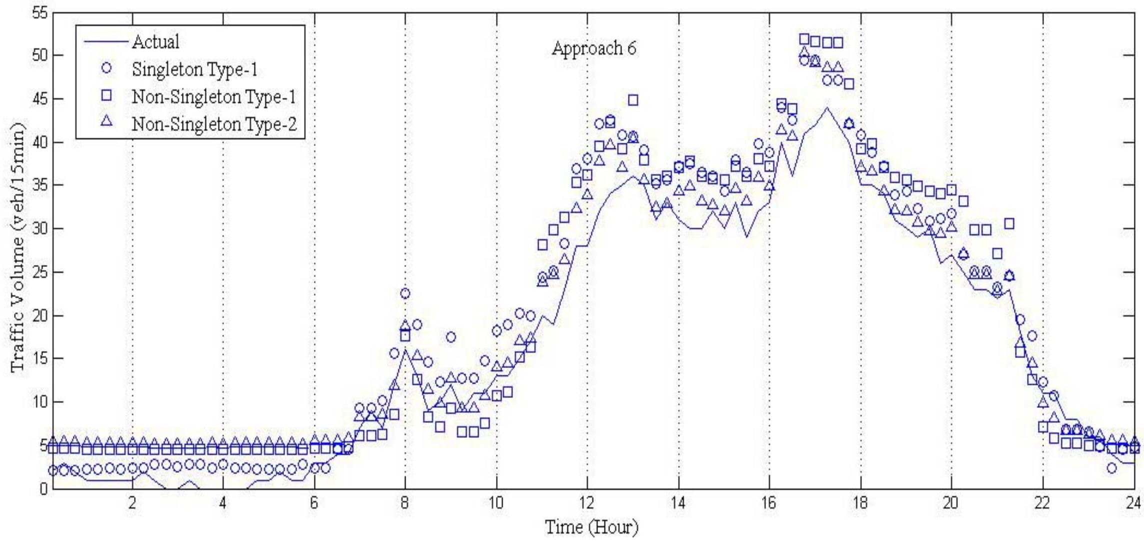


Figure 4-19: The predicted output of singleton and non-singleton type-1 FLS and non-singleton type-2 FLS, and the actual value for the Approach 6 at intersection A.

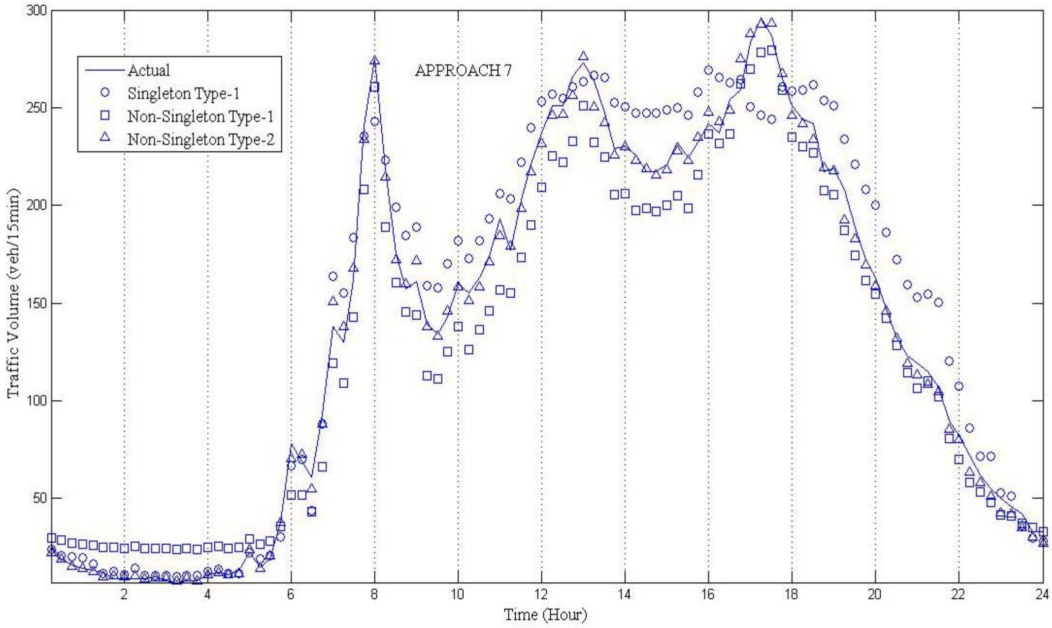


Figure 4-20: The predicted output of singleton and non-singleton type-1 FLS and non-singleton type-2 FLS, and the actual value for the Approach 7 at intersection A.

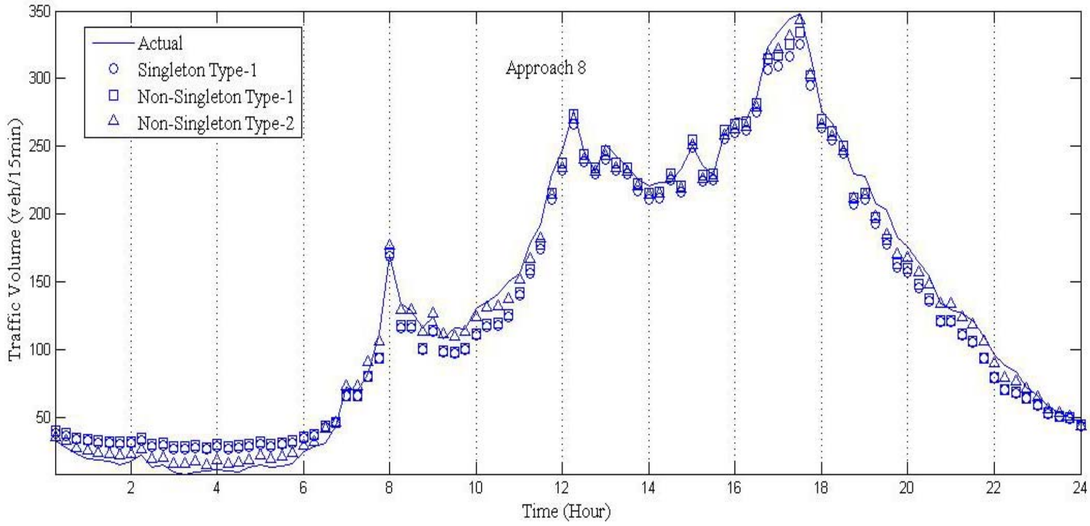


Figure 4-21: The predicted output of singleton and non-singleton type-1 FLS and non-singleton type-2 FLS, and the actual value for the Approach 8 at intersection A.

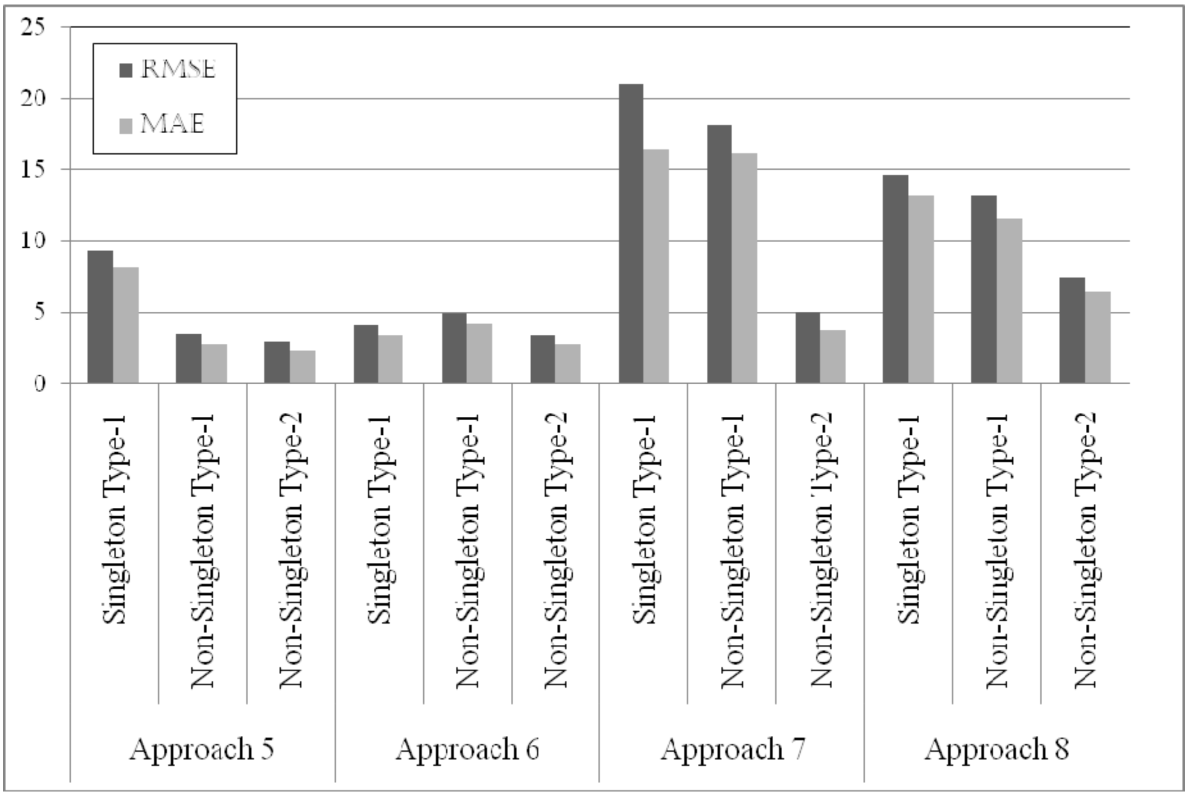


Figure 4-22 : Comparison of prediction accuracy for different models of the approaches of the intersection A.

Table 4-6: Comparison of prediction accuracy for different models of the approaches of the intersection A.

		RMSE	MAE	R-SQ	Mean	Std. Dev.
Approach 5	Singleton Type-1	9.3378	8.1781	0.1513	-8.1781	4.5306
	Non-Singleton Type-1	3.4803	2.7641	0.8821	-2.7558	2.1367
	Non-Singleton Type-2	2.9450	2.3264	0.9156	-2.0787	2.0971
Approach 6	Singleton Type-1	4.1959	3.4129	0.9080	-3.2600	2.6554
	Non-Singleton Type-1	5.0209	4.3224	0.8683	-3.3105	3.7947
	Non-Singleton Type-2	3.3691	2.8166	0.9407	-2.5883	2.1680
Approach 7	Singleton Type-1	21.0278	16.5061	0.9505	-11.4801	17.7099
	Non-Singleton Type-1	18.2109	16.1828	0.9629	10.1155	15.2226
	Non-Singleton Type-2	5.0562	3.8317	0.9971	1.4543	4.8679
Approach 8	Singleton Type-1	14.6953	13.2954	0.9785	5.6768	13.6257
	Non-Singleton Type-1	13.2002	11.6311	0.9827	2.8630	12.9557
	Non-Singleton Type-2	7.5390	6.4475	0.9943	2.2642	7.2287

4.3.9 Conclusion

Type-2 FLS has been successfully applied in wide range of areas which inspired to investigate the potential of this model in intersection traffic flow prediction. The type-2 FLS has more design degrees of freedom associated compared to a type-1 FLS which indicates higher potential of better performance. This study proposed type-2 fuzzy model which was initialized by a type-1 FLS to predict intersection traffic flow and compared the MAE, RMSE and R-square values of type-2 fuzzy models with type-1 singleton and non-singleton models. The MAE and RMSE are higher for the approach 7 and 8 due to the higher traffic flow compared to other approaches. The obtained empirical results of type-2 fuzzy models yield more accurate intersection traffic flow prediction than the type-1 singleton and non-singleton models. The adoption of SVD as a tool to reduce the number of rules and complexity of the model seems promising. It also contributed in reducing the processing time required for training the model. Therefore, this study indicates the potential of SVD based type-2 fuzzy models as a valid alternative for predicting intersection traffic flows.

4.3.10 Summary of Section 4.3

This chapter proposed a SVD based type-2 fuzzy models to predict intersection traffic flow. The adoption of SVD ensures only important rules in the rule base and saves training time to build the model without any significant loss of accuracy. The obtained empirical results of type-2 fuzzy models yield more accurate intersection traffic flow forecasting than singleton and non-singleton type-1 fuzzy logic models. Therefore, this study indicates the potential of SVD based type-2 fuzzy models as a valid and promising alternative for predicting intersection traffic flows.

4.4 Summary of Chapter 4

In this chapter, three AI based models namely AIM model, ANFIS model and type-2 fuzzy logic model are investigated to predict intersection traffic flow. Based on the performance it seems that FCM based ANFIS model outperforms type-2 FL model, type-1 singleton and non-singleton model, SC based ANFIS model, AIM and NN model for the prediction of the traffic flow of Approach 5 and 6. Type-2 FL model performs better than other considered models in predicting Approach 7 and 8. The prediction accuracy of the AIM model is close to that of FCM based ANFIS model and type-2 FL model. It seems that all the considered models are valid and promising alternative for predicting traffic flow.

The considered AIM model is a deterministic and self-organizing AI model. The model building process is quite simple and it doesn't require much intervention from the user. These features of this model seem promising for the transportation engineering practitioners. Another feature of this model is that it can determine important input variables through a built-in process based PSE. This characteristic can be exploited to use the model as input selection model for other AI models in order to predict traffic flow of a very big and complex road network.

The FCM based ANFIS model is not adopted in the literature for traffic flow prediction problem as per the knowledge of the author. This model exploits the capabilities of both neural network and fuzzy logic system like any other neuro-fuzzy model. This model performs better than widely used SC based ANFIS model. The model building process of

FCM based ANFIS model is more complex compared to AIM model and it requires more involvement of the user. Typically, it requires more processing time.

Type-2 FLS has been successfully applied in wide range of areas which inspired to investigate the potential of this model in intersection traffic flow prediction. The type-2 FLS has more design degrees of freedom associated compared to a type-1 FLS which indicates higher potential of better performance. The proposed SVD based type-2 model performs better than type-1 model. The introduction of SVD reduces number of rules and complexity of the model. Type-2 FLS can handle the uncertainty associated with the input data and the linguistic variables. The model building process of type-2 FLS is more complex compared to AIM model and it requires more involvement of the user. Typically, it requires more processing time even higher than FCM based ANFIS model.

CHAPTER 5 TOD BREAKPOINTS DETERMINATION AND EVALUATION

This chapter is intended to investigate the usefulness of the predicted traffic flow data for determining TOD breakpoints instead of the observed data. The methodology which is mentioned in Chapter 3 is used to determine TOD breakpoints using the observed 24-hours traffic flow of a month. In Chapter 4, different models were proposed to predict the traffic flow for the same month. Now, the best models for each approach of Intersection A are selected based on their performance. The ANFIS-FCM models for Approach 5 and 6 outperform other considered models. The type-2 FL models for Approach 7 and 8 perform better than other models. The best models are used to determine TOD breakpoints. The obtained breakpoints are evaluated for the observed traffic flow by using SimTraffic, microscopic simulation software. The TOD breakpoints obtained using original observed traffic flows are also evaluated in the simulation environment. Then the MOEs are compared to investigate the effectiveness and efficiency of using predicted data rather than the original observed data.

The present study emphasizes on the statistical clustering approach because it is easier to implement for the practitioners especially in developing countries such as Saudi Arabia. Specifically, this research attempted to improve the statistical approaches adopted by Smith et al (2001) and Wang et al (2005) which are dependent on the judgment of the users.

5.1 Clustering Algorithms

K-means is one of the most popular data mining and unsupervised learning algorithms that solve the well known clustering problem which follows a simple and easy way to classify a given data set through a pre-specified number of clusters K (Chiang and Mirkin, 2007). The K-means algorithm starts with an initial partition and centers then each entity is assigned to the cluster having the closest center. When all points have been assigned to one cluster, the positions of the centers are reordered. Finally, the last two steps are repeated until cluster membership does not change. However, the final clustering also depends on the initial partition like many other types of numerical minimization. Due to the starting partition it might happen for K-means to reach a local minimum, where reassigning any one point to a new cluster would increase the total sum of point-to-centroid distances although there exists a better solution.

In order to provide the initial solutions for K-means, other clustering techniques such as subtractive clustering method can be used. Chiu (1994) developed subtractive clustering which is a modified form of the mountain method for cluster estimation. In this method, the potential of each data point to become a cluster center is calculated depending on the density of surrounding data points. Consider *m* data points $(X_1, X_2, X_3, \dots, X_m)$ with *n* dimensions which are assumed to have fallen inside a hyper box. The density (ρ_i) of the data point X_i can be expressed as:

$$\rho_i = \sum_{k=1}^{m} exp\left(-\frac{\|X_i - X_k\|^2}{\left(\frac{r_a}{2}\right)^2}\right)$$

r_a is a positive number which defines the influence area of a data point and the data points beyond the radius do not have influence on the density of X_i.In the next step, the data point with the highest density value is selected as the first cluster center. After this selection, the density of a data point is changed by the following formula.

$$\rho_i = \rho_i - \rho_t exp\left(-\frac{\|X_i - X_t\|^2}{\left(\frac{r_b}{2}\right)^2} \right)$$

r_b is a positive number which defines the influence area where the function of the density of data point will reduce. $\rho_t$is the density of the first selected cluster center (X_t) among the data points. After changing the density of data point, the new cluster center will be selected and the density of all data points will be amended again. This process will be continued until all of the data point is within radii of a cluster.

5.2 Feature Selection

The earlier hierarchical, and non-hierarchical clustering technique such as K-means based approach to find out optimal TOD breakpoints suffered from "outliers" of clusters (Smith et al, 2001 and Wang et al, 2005) although they considered a list of features such as volume, occupancy etc. In this study, Z-score of all approaches of the Intersection A and time variable are considered as input features for implementing clustering technique. The application of clustering technique using 15-min traffic counts in finding out the number of clusters creates "outliers" because the relative value of the traffic counts is too much dominating compared to the time variable. Therefore, the distance between the points which play the significant role in determining the cluster is mainly influenced by the traffic counts. The use of Z-score instead of the traffic counts ultimately provides greater relative importance to the time variable in determining the cluster.

In order to investigate the effect of features in K-means algorithm the 15-min traffic counts were clustered to determine number of transitions. The traffic counts along with time variable were also used to determine the number of transitions. Finally, the Z-score of traffic counts and time variable are used for the same purpose. The obtained results are shown in Table 5-1. The solution that K-means provides often depends on the initial assignment of the clusters. When the K-means technique was applied, fifty random initial locations of centers were considered in order to get consistent results with respect to minimum sum of squared distances between the items and their cluster centroids. The numbers of transitions obtained by using 15-min traffic counts are not practical (Table 5-2). For each number of clusters at least there is a cluster of TOD which continues only for 15 minutes. These types of TODs are termed as impractical from implementation point of view. If the duration of TOD is short (less than 30-min), then the progression bandwidth obtained through coordination of traffic signals along a street may be disrupted. Depending on the prevailing traffic condition and from practical implementation point of view, it is assumed that the TOD having duration less than 30-min may be considered as not reasonable.

Table 5-1: Number of transitions obtained by using only15-min traffic counts, and Z-score of the traffic counts along with time variable.

No of Clusters	Number of Transitions	
	For 15-min Traffic Count	For Z-score of 15-min Traffic Count and Time Variable
2	6	1
3	6	2
4	8	3
5	10	4
6	11	5
7	21	6
8	18	7
9	15	8
10	19	9

Table 5-2: Transition time obtained by using only 15-min traffic count.

Number of Clusters	2	3	4	5	6	7	8	9	10
Transition Time (Hour)	7.50	6.75	6.75	5.75	5.75	5.75	5.75	5.75	5.75
	8.50	7.75	7.75	7.25	6.75	6.75	6.75	6.75	6.75
	8.75	8.00	8.00	7.75	7.50	7.50	7.50	7.50	7.50
	9.00	11.25	11.25	8.00	7.75	7.75	8.25	8.25	8.25
	9.75	19.50	16.25	11.25	8.00	8.00	8.50	10.25	9.00
	20.50	22.00	17.75	16.25	8.50	8.50	8.75	11.25	9.25
			19.50	17.75	10.25	8.75	9.00	11.75	9.50
			22.00	19.50	11.50	9.00	9.75	13.50	10.75
				21.50	16.50	9.75	11.25	15.50	11.50
				23.50	17.75	11.25	11.75	16.50	11.75
					19.25	11.75	13.50	17.75	13.50
					20.25	13.50	15.50	18.50	15.50
					21.75	14.75	16.50	19.50	16.50
						15.00	17.75	20.50	17.75
						15.50	18.50	22.00	18.50
						16.50	19.50		19.25
						17.75	20.50		20.00
						18.50	21.75		20.50
						19.50			22.00
						20.50			
						21.75			

5.3 TOD Breakpoints Determination

In this study, it is attempted to determine number of clusters among the data points and the location of TOD breakpoints. If there is no clear idea about the number of clusters for a given set of data, subtractive algorithm is a fast, one-pass algorithm for estimating the number of clusters and the cluster centers (The Mathworks, 2010). The cluster estimates, which are obtained from this algorithm, can be used to initialize iterative clustering methods such as K-means technique.

Subtractive clustering method is used to determine the initial cluster centers and the natural number of clusters considering the four approach volumes as the input along with a radius of 0.3 to 0.5. Based on the experiments, the radius is fixed to 0.5. The obtained cluster centers are provided as the initial solutions for the K-means technique to determine TOD breakpoints. Subtractive algorithm provides five numbers of clusters for the original observed data and four numbers of clusters for the predicted data. Finally, K-means determines the TOD break points which are shown in Figure 5-1 and Figure 5-2. In order to implement K-means technique it is required to provide number of clusters and optionally, the initial cluster centers can also be provided. In this study, both the number of clusters and the initial cluster centers are provided. This technique is implemented in the MATLAB™ environment. The output of K-means technique provides the cluster number of all the input. In this case, each input consists of four approach flows of the middle intersection (A). In Figure 5-1, the first cluster starts at 00:00 hour and ends at 05:45 hour. It means the first TOD will continue for the same period and the breakpoint is 05:45 hour. The investigation of Figure 5-1 and Figure 5-2 reveals that the TOD breakpoints are close for observed and predicted data. Moreover, the duration of TOD for both cases is free from “unclean” clusters and the obtained TOD can be used to develop traffic signal timing.

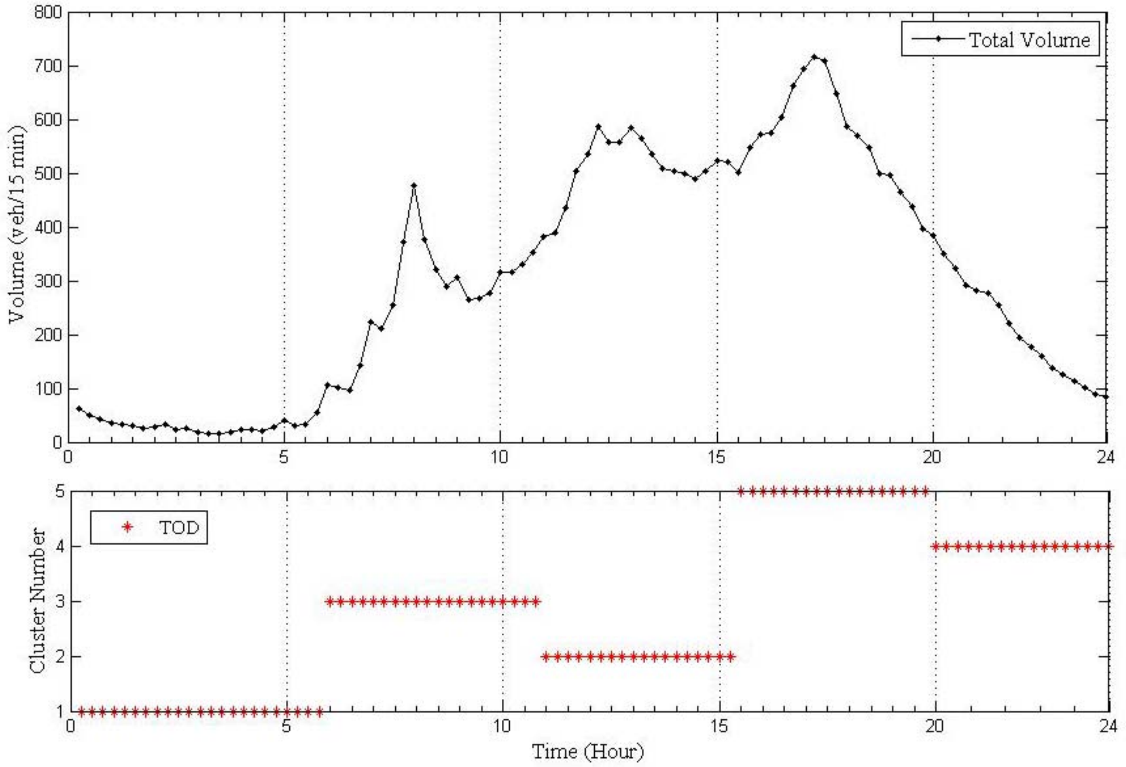


Figure 5-1: The break points of time of day using subtractive method based K-means clustering technique for the observed four approach traffic flows of the middle intersection (A).

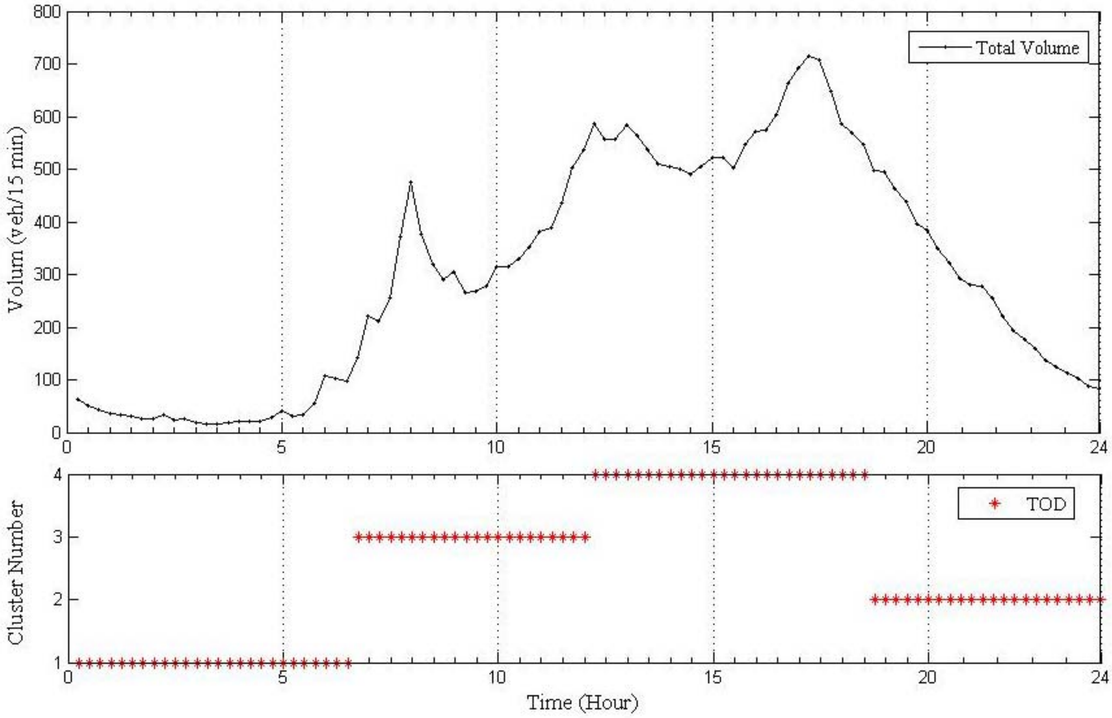


Figure 5-2: The break points of time of day using subtractive method based K-means clustering technique for the best predicted four approach traffic flows of the middle intersection A.

5.4 Evaluation of TOD Breakpoints through Simulation

In order to investigate the usefulness of the predicted traffic flow data for determining TOD breakpoints instead of the observed data, SimTraffic is used to compare the values of the measure of effectiveness (MOE)s for each case. In the evaluation process in SimTraffic, the original observed traffic counts are used as described in the Chapter 3. In order to make the obtained MOEs more reliable it is recommended to use multiple microscopic simulation runs in SimTraffic by adopting random seeds. Total delay, total stops and fuel used (gal) are considered as MOEs. In the simulated environment, the effect of the transition of traffic signal timing plans at a TOD breakpoint is also considered to get realistic values of the selected MOEs. By adding the values of MOEs for 24 hours the performance of TOD breakpoints are evaluated for both the observed and predicted traffic counts.

Table 5-3 shows that the traffic signal timing plans for the TODs obtained by using predicted traffic counts cause an increase in the values of MOEs. The predicted traffic counts cause increase of total delay (hr), total stops and fuel used (gal) are 3.4, 4.4 and 0.7 percent compared to the corresponding values obtained by using observed traffic counts. “Total delay is equal to the total travel time minus the travel time for the vehicle with no other vehicles or traffic control devices” (Trafficware, 2006). It also include all time spent by the vehicles which are denied to enter while they are waiting to enter the network. The total stops are obtained by adding a stop whenever a vehicle’s speed drops below 10ft/s. When the speed of a vehicle reaches 15ft/s, then it is considered going again. The fuel used is determined by the vehicle’s fleet, speed and acceleration.

The MOEs are reported in Table 5-3. The higher values of MOEs indicate increased disutility and lower values mean decreased disutility. It is aimed to determine the TODs and design the traffic signal timing plan in such a way that it will reduce the values of MOEs. If the TODs and traffic signal timing plans are obtained by using the predicted traffic flows then it may not work as efficiently as the TODs and timing plans obtained by using actual traffic flows. It also happened in this study and it causes increase in the values of MOEs but the raise is not significant. It indicates the potentiality of using predicted traffic flows instead of actual traffic flows to determine TODs and the corresponding traffic signal timing plans.

Table 5-3: Performance of traffic signal timing plans for the corresponding TOD breakpoints obtained by using observed and predicted traffic counts.

MOEs	Observed Traffic Count	Predicted Traffic Count	Increase (%)
Total Delay (hr)	204.4	211.4	3.4
Total Stops	21230	21274	4.4
Fuel Used (gal)	1506.0	1516.7	0.7

5.5 TOD Breakpoint Determination: Alternative Approach

A sequence of observed data which are ordered in time can be defined as time-series. This kind of data can be obtained from business, science, engineering etc. It is a data mining problem to partition a time-series into internally homogenous segments (Abonyi et al., 2005). Typically, the segmentation aims to locate stable periods of time, identify change points, or simply compress the original time-series (Last et al., 2000). Most of the available segmentation algorithms focus on only one time-variant variable although many variables are simultaneously tracked and monitored in many real-life applications (Kivikunnas, 1998).

The traditional clustering algorithms are intended for independent observations which assumption is violated for time-series datasets (Beaver and Palazoglu, 2007). Those algorithms cluster observations based on distances, yielding clusters of distinguishable means but dynamic events do not have a constant mean, and thus cannot be properly detected (Beaver and Palazoglu, 2007). Beaver and Palazoglu (2007) proposed a PCA based clustering algorithm for clustering auto-correlated and cyclic datasets of chemical processes. In their proposed framework, separate cluster analyses which were performed at different time scales were combined to identify all process states and accurately determine the transitions points.

In the first attempt of this study, clustering is performed by considering time variable and Z-score of all the approaches of the middle intersection. Abonyi et al. (2005) proposed a PCA and fuzzy sets based approach to segment multivariate time-series in which time variable was considered. Park et al. (2003a) commented on the “unclean” clusters that these are inevitable since statistical clustering algorithm do not consider time variable. But the author of this study didn't find any attempt in the literature to use time variable to solve this problem.

The obtained result by using the mentioned features solves the problem of "unclean" clusters. But this approach will force the K-means not to use the early period data of a day for clustering with the late night traffic flow data. In reality, these traffic flows are occurring consecutively. It indicates that the K-means is being used to cluster linear data but practically the data have cyclic characteristic. Apart from the cyclic characteristic, this problem can be defined as constrained clustering in which all data points of a cluster should come from successive time points.

In order to solve the above mentioned problem, the monthly average of 15-min traffic count of an average day is considered twice in using the subtractive clustering based K-means approach. This clustering effort provides the cluster which will start at late night and continue in the next day morning. This obtained cluster is fixed and the same clustering technique is used for the remaining parts of the day. Therefore, the fixed cluster (TOD) starts at 20:30 pm ends at 06:15 am (Figure 5-3). The features of the remaining part between 06:30am to 20:30pm are used for the subtractive clustering based K-means technique. The obtained TODs are shown in Figure 5-4.

The provided alternative approach theoretically takes care of cyclic characteristics of the traffic flow data. It can be assumed that this novel approach will ensure better performance compared to the approach which is illustrated before in this Chapter.

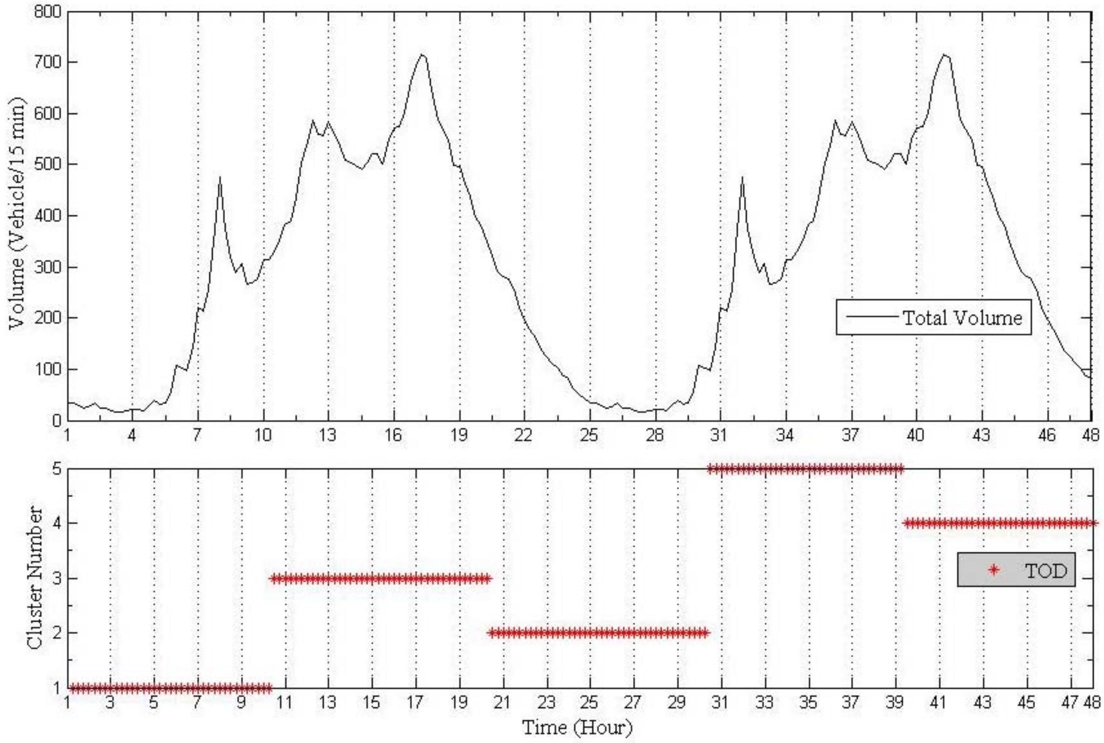


Figure 5-3: TOD breakpoint obtained by using same monthly average data two times.

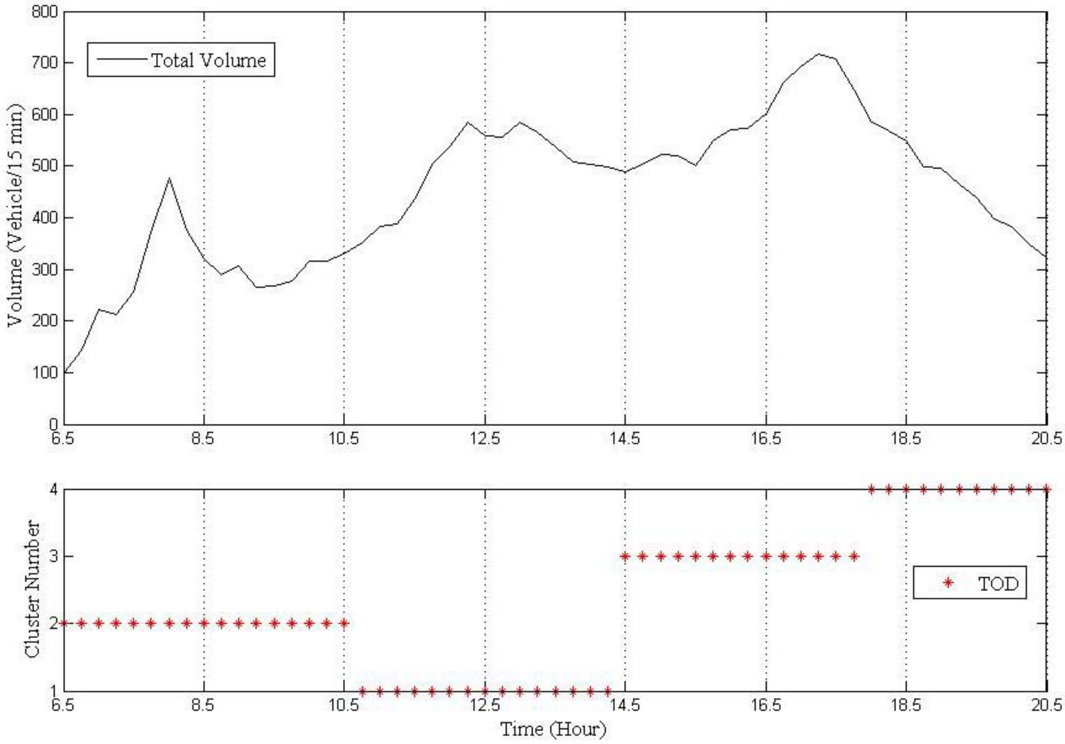


Figure 5-4: The TOD breakpoint obtained for the period between 06:30 am and 20:30 pm.

5.6 Conclusion

This study solves the problem of "unclean" clusters by selecting appropriate features which include Z-score of the traffic count and the time variable. Typically, the TOD breakpoints are obtained through judgmental approach. The proposed approach is systematic in nature and it does not require any subjective intervention. This chapter used the predicted data instead of observed data for determining TOD breakpoints. The use of predicted traffic counts to determine TODs and the corresponding traffic signal timing plans causes increase of MOEs compared to the corresponding values obtained by using observed traffic counts. It indicates the potentiality of using predicted traffic flows instead of actual traffic flows to determine TODs and the corresponding traffic signal timing plans. In this chapter a novel alternative approach is also studied to consider the cyclic characteristics of traffic flow data in the process of determining TODs.

5.7 Summary of Chapter 5

The practitioners typically use judgmental approach to determine TOD breakpoints for both pre-timed and actuated traffic signal controllers. In order to determine the minimum and maximum green time for actuated controllers, it is required to adopt TOD mode. On the other hand, the statistical clustering approaches attempt to solve the problem but most of the proposed approaches require some sort of intervention. This chapter proposes time variable and Z-score of the approach traffic counts as the prospective features for determining TOD breakpoints. The obtained results are promising as it solves the problem of "unclean" clusters.

This chapter also used the predicted traffic flows instead of the observed traffic flows to determine TOD breakpoints by using K-means clustering technique. The initial cluster centers and number of clusters used in the K-means is obtained by adopting subtractive clustering algorithm. The subtractive clustering method determines the cluster center considering the traffic flow data. The subtractive algorithm based K-means determines the TOD breakpoints for predicted and observed data. The TOD breakpoints are evaluated in SimTraffic using observed data. The use of predicted traffic counts causes increase of total delay (hr), total stops and fuel used (gal) by 3.4, 4.4 and 0.7 percent compared to the corresponding values obtained by using observed traffic counts. The results indicate the potentiality of the use of the predicted data for TOD breakpoint determination and other similar kind of problems.

CHAPTER 6 PREDICTION MODELS FOR LOCAL DATA

In order to realize the regional traffic characteristics, generally the transportation authority operates permanent traffic counting sites to measure traffic volume along regional highway network. But there is no systematic methodology reported in the literature to select the sites depending on the available datasets in Saudi Arabia. In this study, it is attempted to predict the annual average hourly variations of traffic at a few freeway stations with the help of the data of other stations. If the data of a few locations can be predicted accurately based on the data of other locations for a certain years, then those stations can be left without collecting data for the future year based on the assumption that the developed model will perform same for the future year.

If the transportation authority can develop reliable prediction model for freeway traffic flow then they can find out the appropriate locations which should be considered for traffic flow data collection. It will help the transportation authority to invest the saved money in collecting the data of other required locations. This study attempted to develop freeway data collection strategy. Figure 6-1 shows the proposed traffic data collected strategy for Ministry of Transportation (MOT) based on the prediction models used in this study. According to this strategy, MOT will prepare the datasets by compiling the traffic flow data of all the freeway stations. Then the prediction models will be developed for each station using the data of all other remaining stations. Then the models of the stations will be ranked based on the performance measurement. Depending on the importance and rank of the models, MOT can select the freeway sites from which traffic flow data should be collected. The problem of model building for freeway traffic flow is challenging and different from intersection traffic flow prediction. Moreover, the considered freeway sites are not in located in the closed vicinity rather those are dispersed significantly. It makes the input selection procedure for the model very difficult.

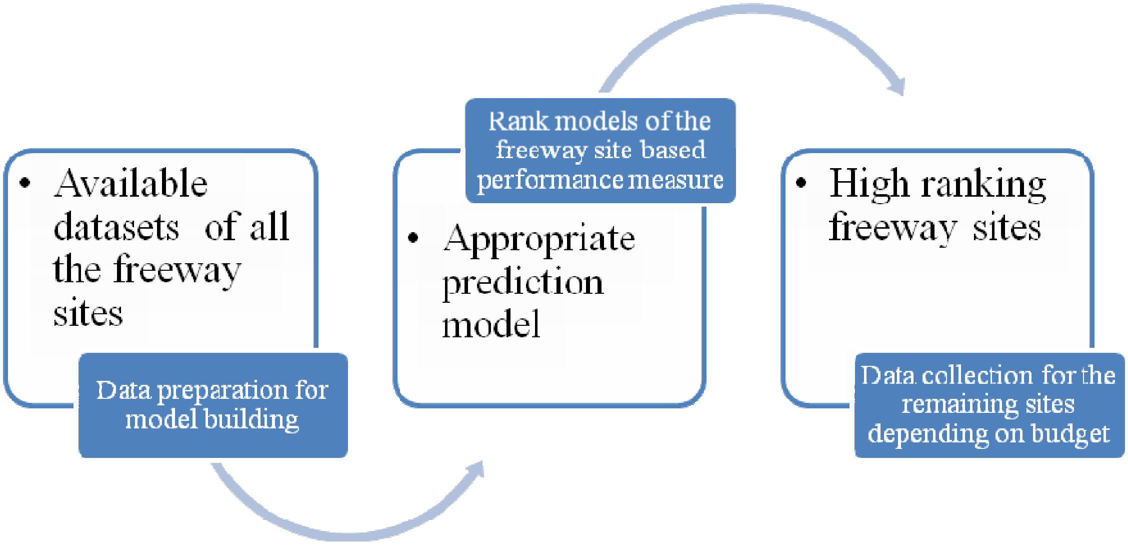


Figure 6-1: Proposed traffic count data collection strategy for MOT based on prediction models.

Any kind of transportation study relies heavily on traffic count data. The traffic count data are used in many traffic studies, which include highway geometry, benefit of roadway improvements, estimates of road revenue, selection of highway routes, selecting the timing of maintenance, signal timing, location and design of highway systems, air quality analysis, design of traffic control systems and accident rates, average daily traffic, and location of service areas (Federal Highway Administration, 2001). With the help of continuous volume data, intersection operation and safety can be evaluated. It can also help in evaluating growth patterns and in calibrating transportation planning models. Even businesses can determine the amount of exposure of a store or advertising sign using the traffic volume data. Although traffic count data plays an important role in many traffic studies, the associated cost is also of concern. The total costs for a roadside detector include capital costs (purchase and installation) and operational costs (maintenance, support and day-to-day operation). The capital cost of a detector ranges from $700 to $29,000, operation and maintenance cost varies from $100 to $2300, and the lifetime of the detector varies from 5 to 20 years (US DoT, 2007).

In this study, a hybrid model structure is proposed to predict freeway traffic volume in which the important inputs are selected by using abductive network. Then the selected inputs are used to build ANFIS model, type-2 fuzzy logic model and ANN model. Finally, the best performing model is selected.

6.1 Regional Road Network System of Saudi Arabia

The road network plays a significant role in the Saudi economy by contributing in the development of some sectors including minerals, agriculture, industry and commerce

(Aldagheiri, 2009). A properly developed road network reduces the cost of transportation both in terms of time and money, and integrates various regions within the country (MOT, 1971). It also helps in understanding and interacting with the neighboring countries. Generally, in Saudi Arabia, most of the development projects, whether for public services, religious purposes, agriculture or industry, rely heavily on the construction of new roads (MOT, 1985).

It is reported that the first modern road in Saudi Arabia was built in 1938 by Egypt which connected the city of Jeddah with the holy places in Makkah and was mainly intended to serve the pilgrims arriving at the seaport of Jeddah (Al Sayyad, 1956). Before the first development plan the total length of paved road was close to 8,000 kilometers and 20 percent of which was built by Arabian-American Oil Company (MOT, 1986). The modern network of roads in the Kingdom of Saudi Arabia has been constructed over a period of three decades which has become a symbol of the modernization of the country and a national land mark (Aldagheiri, 2009).

During the first development plan (1970 to 1975), 3,221 kilometers of new roads were constructed which included 1800 kilometer long road connecting northern and southern part of the country (MOP, 1975). During the second development plan (1975-1980), 11,399 kilometers of paved main roads were constructed which included the magnificent viaducts to access the mountainous south-west part of the country (MOP, 1980). The third development plan (1980 to 1985) focused on inter-city roads due to the increased traffic caused by industrial and agricultural development. During the third development plan, many inter-city roads were upgraded to expressways such as Riyadh-Dammam (388 kilometers), Riyadh-Sedir-Qassim (353 kilometers), Jeddah-Makkah (70 kilometers) and

Makkah-Madinah (424 kilometers) making the total road network to 25,000 kilometers of paved road (MOP, 1985).

The fourth development plan (1985-1990) emphasized on road maintenance along with the new construction of secondary and feeder roads, as well as on links to new and expanding agricultural areas. During this period, the total length of roads increased to 81,500 kilometers including 3,500 kilometers of expressways and divided highways (MOP, 1990). A road of 1,275 kilometers long was constructed during the fourth development plan connecting Dammam, Riyadh and Jeddah.

During the fifth development plan (1990-1995), 4,322 kilometers of paved roads was built (MOP, 1997) and by 1991 the country achieved a road network system that linked all the populated areas of the Kingdom (Aldagheiri, 2009). During the sixth development plan (1995-2000), the length of paved road network reached 45,500 kilometers (MOP, 2000). At the end of this plan, almost all of the towns and cities and villages of the Kingdom were connected by at least two-lane roads (Aldagheiri, 2009). During the seventh development plan (2000-2005), the length of paved road network increased to 51,800 kilometers (MOT, 2003) and a 810 kilometer long highway was completed connecting Al-Qassim, Madinah, Yanbu, Rabigh and Thuwal. In 2007, it is reported that the regional highway network of Saudi Arabia comprised of 4,621 km expressways, 5,746 km dual lane highways, and 35,575 km single lane highways (MOT, 2007).

6.2 Data Source and Description

In order to realize the traffic characteristics, MOT operates permanent traffic counting sites to measure traffic volume along regional highway network throughout the Kingdom

(MOT, 2007). The counting sites are located outside urban areas as shown in the figure. Therefore, traffic volumes at the selected sites reflect mainly inter-urban traffic.

This study gathered and compiled average hourly variations of traffic over the year of a few stations which are monitored by MOT. The sources of those data are the annual traffic data reports provided by MOT. The interested readers can consult MOT, 2006; MOT, 2007; MOT, 2008. In this study, eleven traffic counting sites are selected to predict the annual average hourly variations of traffic in four stations. The location of the sites is shown in Figure 6-2. The prediction of traffic flow is made for the year 2008. Among the complied data only four stations have the average hourly variations of traffic flow data for the year 2008.

A few statistical measures of the data are reported in the Table 6-1. The mean hourly traffic flow varies between 90 and 2500. The standard deviations of the data are high which indicate higher variability of the data. Skewness is a measure of the asymmetry of the data around the sample mean and its negative value indicates that the data are spread out more to the left of the mean than to the right. On the other hand, the positive value of skewness means that the data are spread out more to the right. If the value of skewness is zero, it can be concluded that the distribution is normal distribution or any perfectly symmetric distribution. The skewness values of the data used in this study revealed that majority of the approach traffic flows are spread out more to the right of the mean and there is no clear indication that the data are generated from any perfectly symmetric distribution process. Kurtosis reveals the outlier-prone characteristics of a distribution and the kurtosis of the normal distribution is 3. The kurtosis values of the data indicate that all the traffic approach flows are less outlier-prone than the normal distribution.

The annual average hourly traffic volume data for the year 2008 are available only for four stations among the selected datasets. The data of three of those stations having the IDs 311, 313,109 and 315 are modeled and predicted. The data of other stations are kept in the pool of input datasets. As the MOT selects different stations in different years for collecting data, the data are not continuously available for the stations. In this study, the data of eleven years are considered and the missing data of the selected stations are filled with the help of linear interpolation.

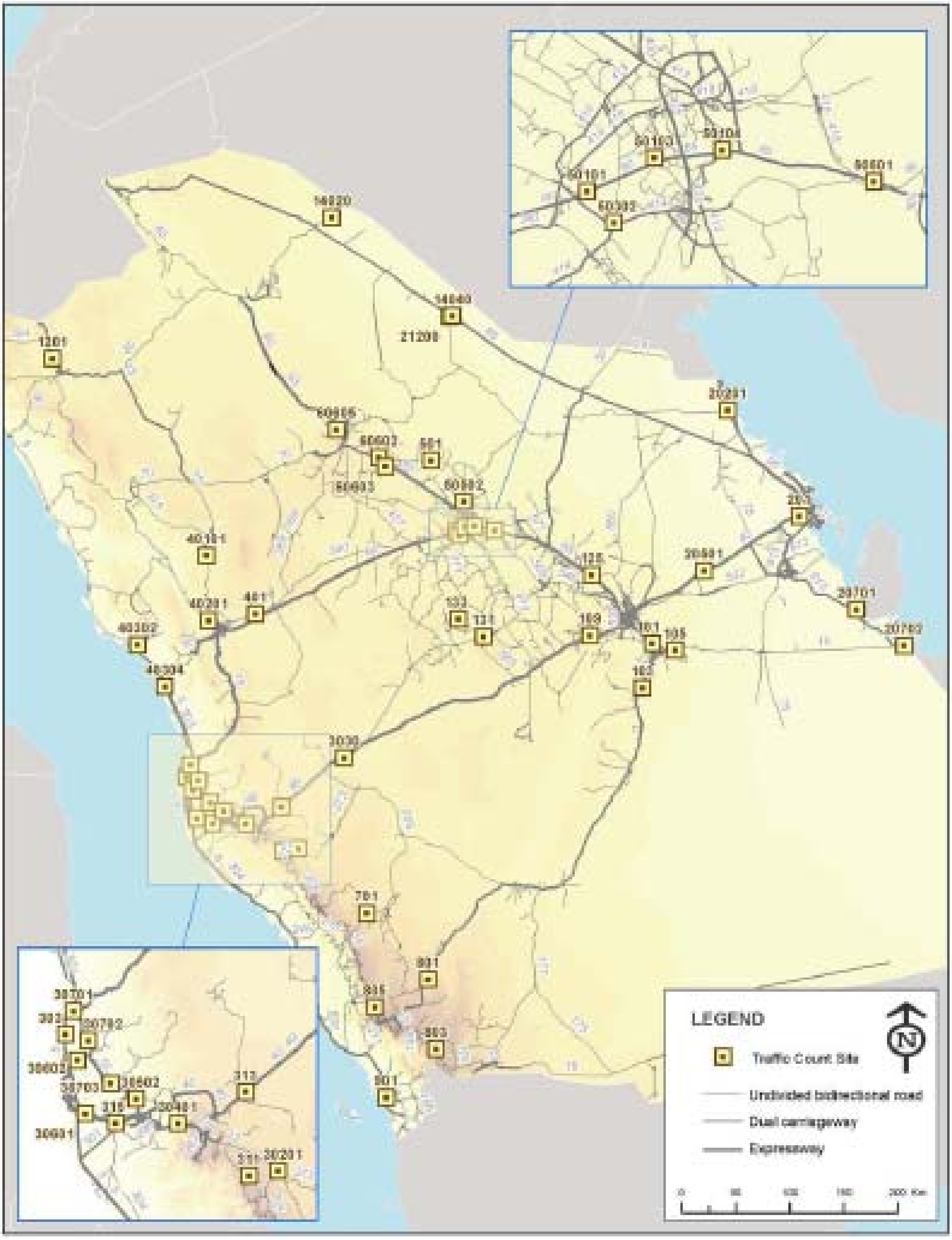


Figure 6-2: MOT traffic count sites (source: MOT, 2007)

Table 6-1: Statistical measures of the annual average hourly traffic flow data of the selected freeway station sites.

Station ID	125	109	303	311	313	315	401	801	803	805	1201
Mean	835	706	776	101	514	3622	133	143	168	437	86
Std Dev	340	298	231	48	146	1067	47	52	65	130	39
Max	1657	1462	1206	213	859	5596	207	410	285	795	146
Min	307	206	340	25	237	1493	40	65	46	225	16
Kurtosis	0.14	-0.18	-0.61	-0.51	-0.18	-0.52	-0.63	10.10	-0.75	1.29	-1.14
Skewness	0.77	0.30	-0.27	0.15	-0.13	-0.58	-0.78	2.11	-0.34	0.80	-0.19
Year of Missing Data	2008	-	2008	2006	2006	2006	2008	-	2008	2006, 2007	2008

6.3 Model Description

In this study, a hybrid model consists of AIM model is proposed to predict the traffic flow data. The model is shown in the Figure 6-3. AIM model is used to select the important inputs. AIM simplifies model development and reduces the learning or development time and effort by automatically determining the optimum model characterized by network size, element types, connectivity, and coefficients (Abdel-Aal, 2004a).

The inputs which are selected by AIM model for building the abductive network are used as input for ANFIS, type-2 fuzzy logic and ANN models as shown in the Figure 6-3. Due to the presence of trend, the datasets were transformed by removing the linear trends. Then ANFIS, type-2 fuzzy logic and ANN models are built. Two ANFIS models are obtained for each station because the initial fuzzy inference systems are developed by using FCM and subtractive clustering methods. The type-2 fuzzy logic model was initiated by a type-1 fuzzy logic system. SVD technique is used to reduce the complexity and number of rules. In order to compare the performance of ANFIS and type-2 FL models a feed-forward ANN is built. Finally, the best performing model is to be selected to predict the traffic volume data of the chosen stations.

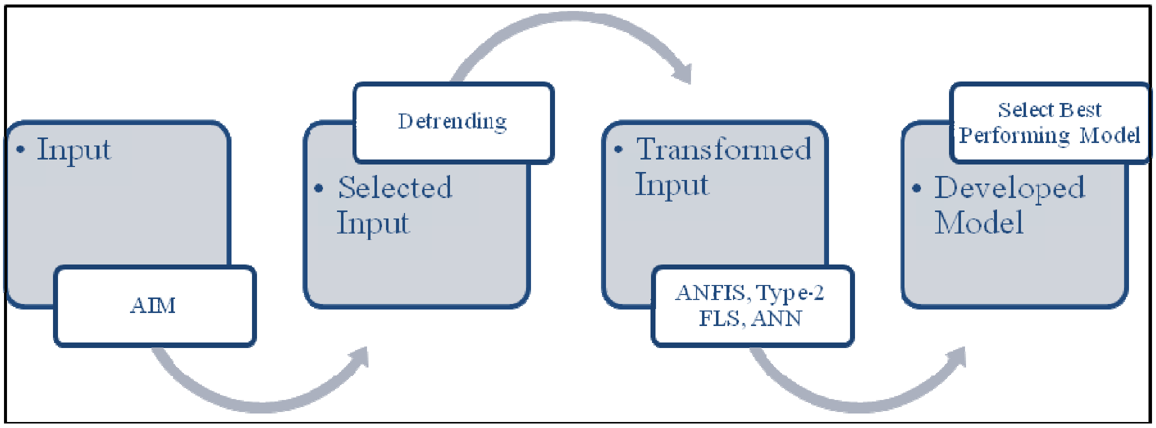


Figure 6-3: Proposed hybrid model for predicting annual average hourly traffic volume data of stations along freeways.

6.4 Model Building

The first step of the model building is to select the input variables. In this study, AIM model is used to select the input. The AIM models are built with the CPM value of 1.00, four numbers of limiting layers and fifteen numbers of inputs for the first layer. These are the default settings of AIM model. The obtained networks and the selected inputs are shown in the Table 6-2 for the prediction model of stations 311, 313,109 and 315. The table indicates that the model of Station 311, Station 313, Station 315, and Station 109 require two, three, four and five number of inputs consecutively. The MAE of the obtained abductive network models for those stations vary between 16 to 540 vehicle/hour. All the models are built by using the default CPM value of 1.00. The size of layer 1 for all the model is selected as ten to provide the model the opportunity to consider all the available inputs.

Table 6-2: Input selection through AIM model.

Output Station (Station ID)	CPM	Input (Station ID)	MAE
311	1.00	109,801	16.287
313	1.00	303,401,801	60.826
315	1.00	125,109,303,803	839.60
109	1.00	303,311,313,401,805	530.95

6.4.1 ANFIS Model

Two ANFIS models are built for each station. The ANFIS models are initiated by two FISs obtained by using fuzzy c-means (FCM) and subtractive clustering (SC) algorithm. In order to implement FCM it is required to provide number of membership functions for input and output. The description of the ANFIS-FCM models is provided in the Table 6-3. For all ANFIS-FCM models, the chosen membership functions are Gaussian functions. The weighted average method is used to defuzzify the output. The models of stations 313, 315, and 109 require 2 numbers of membership functions for each input, and 2 numbers of rules. The model of the Station 313 requires 3 numbers of membership functions for each input and 3 numbers of rules.

The description of the developed ANFIS-SC models is provided in the Table 6-4. It is observed that ANFIS-SC models use more number of membership functions and rules compared to ANFIS-FCM models. The Gaussian function is used as the membership functions for all the inputs. The weighted average method is used to defuzzify the output. The required number of both membership functions for each input and the number of rules for the Stations 311, 313, 315, and 109 are 6,4, 6, and 5 respectively.

Table 6-3: Description of the ANFIS-FCM models.

Model	Number of Membership Functions for Each Input	Number of Rules
311	3	3
313	2	2
315	2	2
109	2	2

Table 6-4: Description of ANFIS-SC models

Model	Number of Membership Functions for Each Input	Number of Rules
311	6	6
313	4	4
315	6	6
109	5	5

6.4.2 Type-2 FLS

There are many methods to automatically construct fuzzy systems by using numerical input-output data. The adopted FLS in this study is shown in Figure 6-3. In this approach all the fuzzy input sets and the random center of consequents are used to build the initial rules which are equal to the sample size. In the next step, the SVD technique is applied to reduce the number of rules and determine the most important rules to produce desired output.

In a type-2 FLS, it is required to decide on the kind of fuzzification (either singleton or non-singleton), select a FOU for membership function, decide functional forms for both the primary and secondary membership functions, and select the initial parameters of the membership function. It is also needed to predefine the type of composition, implication, type-reduction method and defuzzifier. Mendel (2001) stated that there are more design degrees of freedom associated with a type-2 FLS compared to a type-1 FLS which indicates higher potential of a type-2 FLS to outperform a type-1 FLS. In this study, all the proposed non-singleton type-2 models assumed Gaussian primary membership functions with uncertain mean and interval secondary membership functions, product implication and t-norm, center-of-sets type reduction, and defuzzification obtained by the centroid of the type-reduced set.

6.4.3 ANN Model

In order to compare the performance of the ANFIS and type-2 FL model with a typical AI model feed forward neural network is considered. The neural network models are constructed with the same input used in building the corresponding ANFIS and type-2 FL models. The activation functions of the neurons of the hidden layers were “tansigmoid” and the activation function for the output layer was “purelin”. All the input data are normalized in such a way that the mean and standard deviation become “0” and 1 respectively. The description of the model is provided in the Table 6-5. The model of the Station 311 requires only two hidden layers but the models of other stations require 3 hidden layers having 4 neurons in each layer. The topology of the ANN is selected by gradually increasing the number of neurons in each layer and increasing the number of hidden layers. Each model is initialized by random numbers by at least 10 times and then the best performing network is selected.

Table 6-5: Description of ANN models

Model	Number of Hidden Layers	Number of Nodes in Each Hidden Layer
311	2	4
313	3	4
315	3	4
109	3	4

6.5 Performance Evaluation

To measure the performance of the proposed method, mean absolute percentage error (MAPE), mean absolute error (MAE), root mean squared error (RMSE) are reported in the Table 6-6. The predicted values of the models are matched with the actual values to determine the differences. The mean difference (D) and standard deviation (S) of the differences are reported in Table 6-6. The small value of “S” indicates narrow width of confidence interval.

Figure 6-8 shows the performance of the considered models in terms of MAPE. It seems that the MAPE of the best performing models vary between 3 and 12. The MAPE of the Station 313 is the smallest among the all stations.

The type-2 FL model outperforms other models with respect to MAPE, RMSE, and MAE for the prediction model of Station 311. The Table 6-6 shows that the value of “S” of type-2 model is the smallest compared to other models. But the “D” value of ANFIS-FCM model is close to zero with respect to other models. The type-2 FL model fails to perfectly predict the peak traffic flows at 14:00 and18:00 hr although it predicts other points with higher accuracy.

The ANFIS-FCM model outperforms other models with respect to MAPE, RMSE, and MAE for the prediction model of Station 313. The Table 6-6 shows that the value of “S” of ANFIS-FCM model is the smallest compared to other models. But the “D” value of ANFIS-SC model is close to zero with respect to other models. The second best value of

“D” is observed for ANFIS-FCM model. The ANFIS-FCM model predicts the turning points of hourly traffic flow very accurately except at 14:00 hr.

The type-2 FL model outperforms other models with respect to MAPE, RMSE, and MAE for the prediction model of Station 109. The Table 6-6 shows that the value of “S” of type-2 model is the smallest compared to other models. But the “D” value of ANFIS-FCM model is close to zero with respect to other models. The type-2 FL model predicts almost all the turning points with good accuracy except the turning point at 07:00 hr.

The ANFIS-FCM model outperforms other models with respect to MAPE, RMSE, and MAE for the prediction model of station 315. The Table 6-6 shows that the value of “S” of ANFIS-FCM model is the smallest compared to other models. The “D” value of ANFIS-FCM model is close to zero with respect to other models. But ANFIS-FCM model along with other model fail to predict the main peak hour traffic flow at 18:00 hr.

In order to get greater idea about the performance of the models, the recently developed regression error characteristic (REC) is used for further analysis (Bi and Bnnett, 2003; Pina and Zaverucha, 2006; and Torgo, 2005). The REC curve provides a way to visualize and evaluate different regression models which is similar to the receiver operating characteristic (ROC) curves (Fawcett, 2003). The REC curve plots error measures such as absolute deviation or squared residual versus the percentage of points predicted within the tolerance and it provides an estimation of the cumulative distribution function (CDF) of the error. The area-over-curve (AOC) provides a biased estimation of the expected error for a prediction. The interested reader can consult Bi and Bnnett (2003) for further details of REC analysis.

The Figure 6-9,

Figure 6-10,

Figure 6-11, and Figure 6-12 show the REC analysis for the prediction models of site 311, 313, 315 and 109 based on the absolute deviation error. Generally, the figure indicates that REC curves of type-2 FL and ANFIS-FCM models are above the other models which indicate comparative better performance of those models. The REC analysis of the model of the site 311 reveals that around 85% of the predicted values by type-2 FL models are equal or less than the absolute deviation of 20 veh/hr. For the model of the site 313, around 85% of the predicted values by ANN are equal or less than the absolute deviation of 30 but the remaining predicted values extended the absolute deviation up to 59 veh/hr. On the other hand, the ANFIS-FCM model of the site 313 ensures the absolute deviation less than or equal to 40 veh/hr for all predicted data.

The ANFIS-FCM model of the site 315 limits the absolute deviation less than or equal to 400 veh/hr for around 89% of predicted data. The performance of type-2 FL model is also close to the ANFIS-FCM model for same percentage of predicted data but ultimately the absolute deviation increases more than that model. The type-2 FL model of the site 109 ensures the absolute deviation less than or equal to 100 veh/hr for more than 90% predicted data which is far better than other models. But the maximum error of type-2 FL model, ANFIS-FCM and ANFIS-SC are very close to each other.

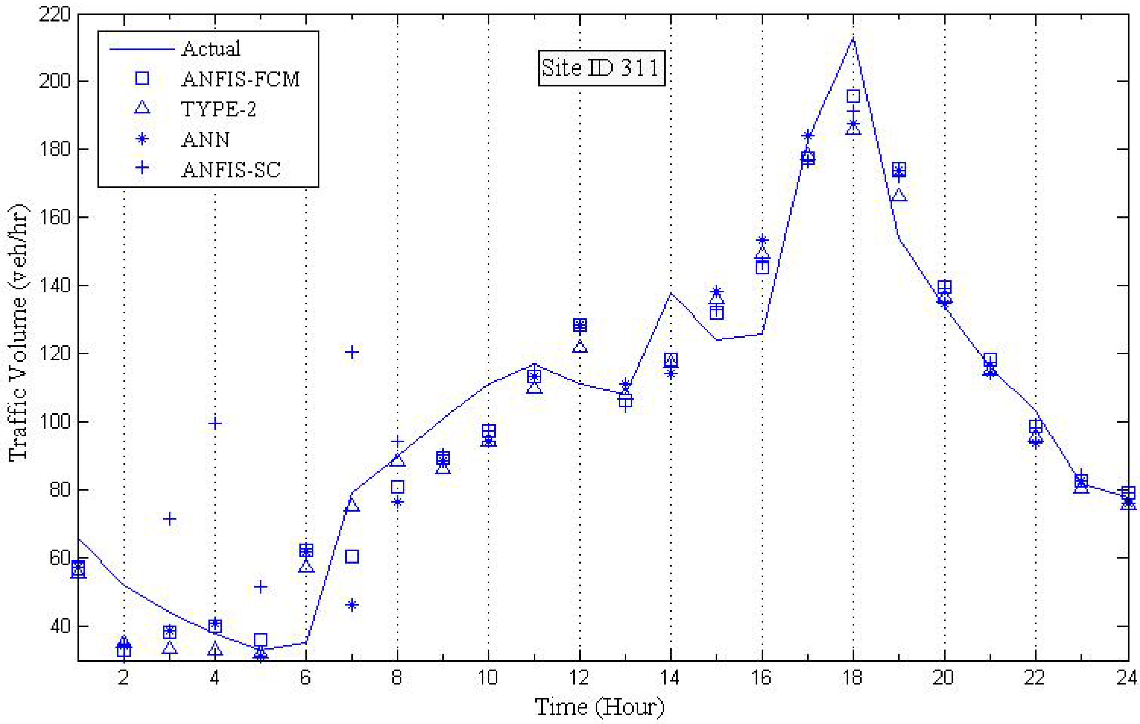


Figure 6-4**:** The predicted and actual values of the models for the freeway site 311.

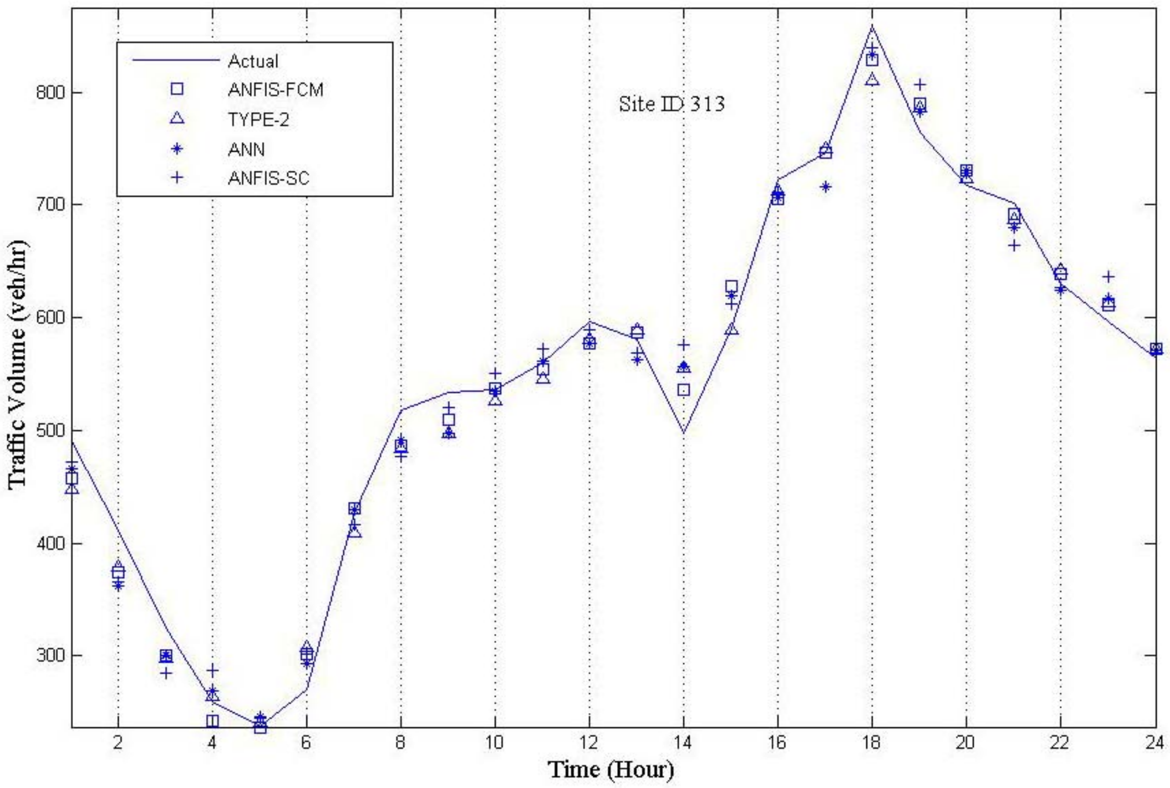


Figure 6-5: The predicted and actual values of the models for the freeway site 313.

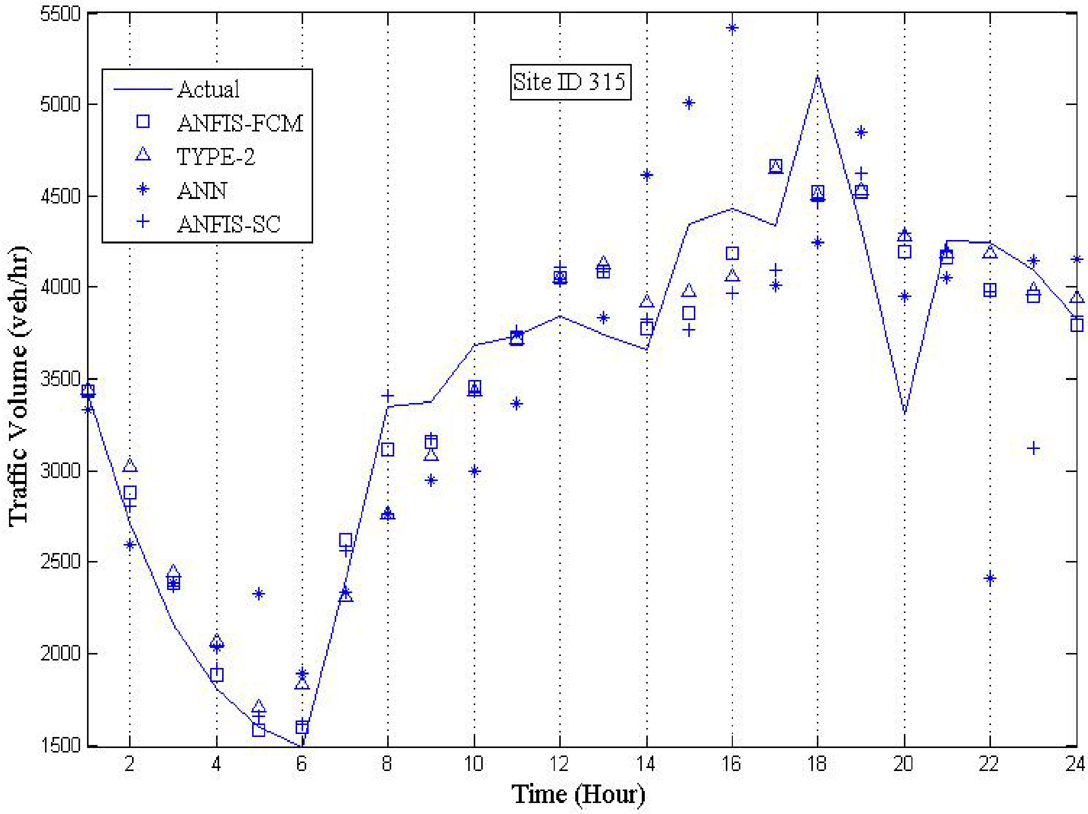


Figure 6-6: The predicted and actual values of the models for the freeway site 315.

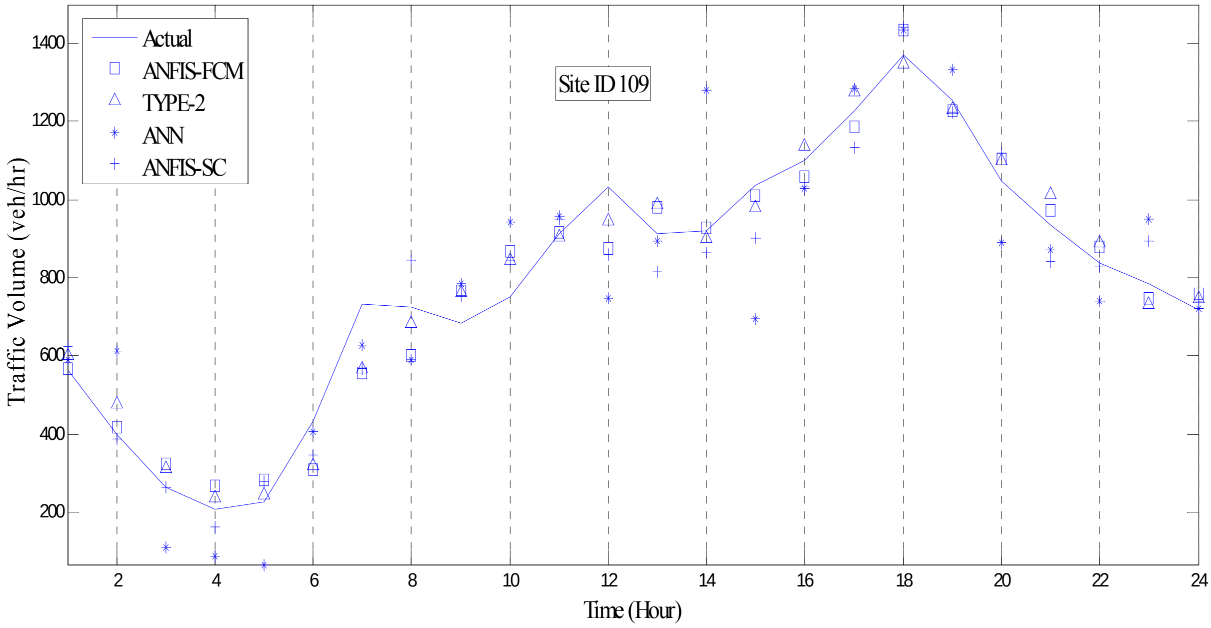


Figure 6-7: The predicted and actual values of the models for the freeway site 109.

Table 6-6: Comparison of forecasting accuracy for different models of the freeway traffic flow prediction.

Station ID	Model	MAPE	D	S	RMSE	MAE
311	ANFIS-SC	24.6474	-6.0593	20.3495	20.8222	15.3686
	ANFIS-FCM	12.7901	1.2926	13.0285	12.8195	10.2788
	ANN	14.3407	2.4761	15.684	15.5521	12.0160
	Type-2 FLS	11.9167	2.8135	12.5276	12.5825	9.8210
313	ANFIS-SC	5.1473	-0.4968	30.711	30.0685	24.6685
	ANFIS-FCM	3.8148	2.5975	22.7183	22.3911	18.5074
	ANN	4.1462	4.3989	24.9407	24.8087	20.5739
	Type-2 FLS	4.1858	5.0094	25.5239	25.4838	20.3499
109	ANFIS-SC	10.2818	14.7781	87.2398	86.6722	74.0238
	ANFIS-FCM	10.1985	1.485	78.1109	76.4807	61.4335
	ANN	21.1105	17.5356	160.478	158.0748	126.3569
	Type-2 FLS	8.878	-7.4936	67.0852	66.0989	56.386
315	ANFIS-SC	7.7366	34.5308	393.2608	386.5262	277.4439
	ANFIS-FCM	6.5087	-11.6889	308.9122	302.6338	228.1548
	ANN	14.607	-17.6263	642.1933	628.919	486.2167
	Type-2 FLS	9.7258	-96.9044	359.678	365.1963	293.1183

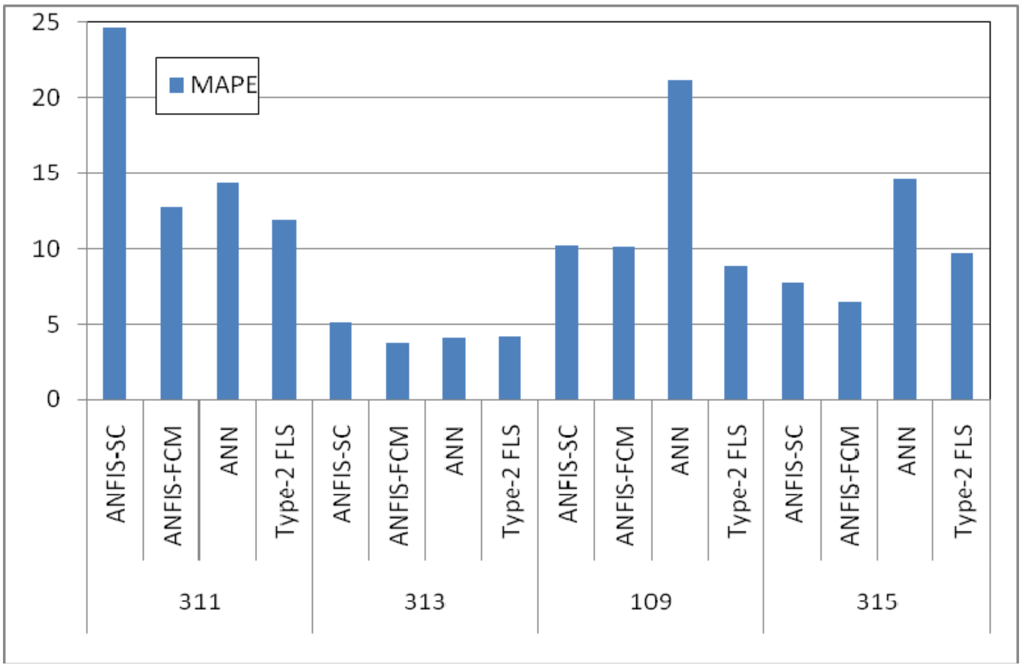


Figure 6-8: Comparison of forecasting accuracy (MAPE) for different models of the freeway traffic flow prediction

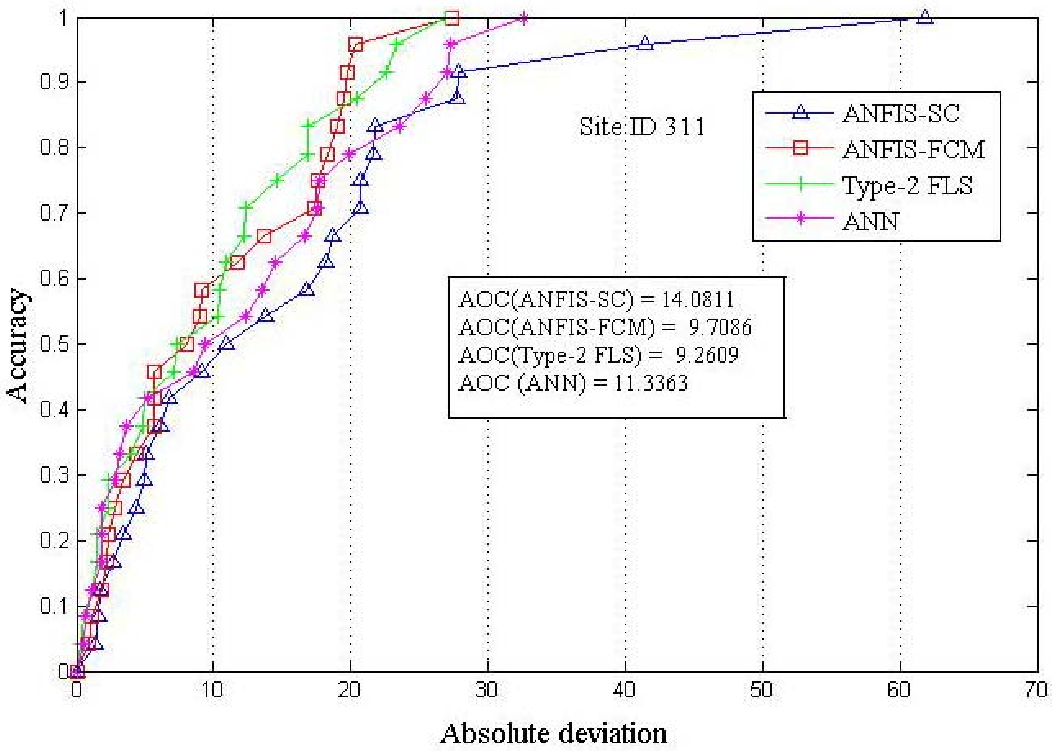


Figure 6-9: REC analysis of different models of the freeway traffic flow prediction at the Station 311.

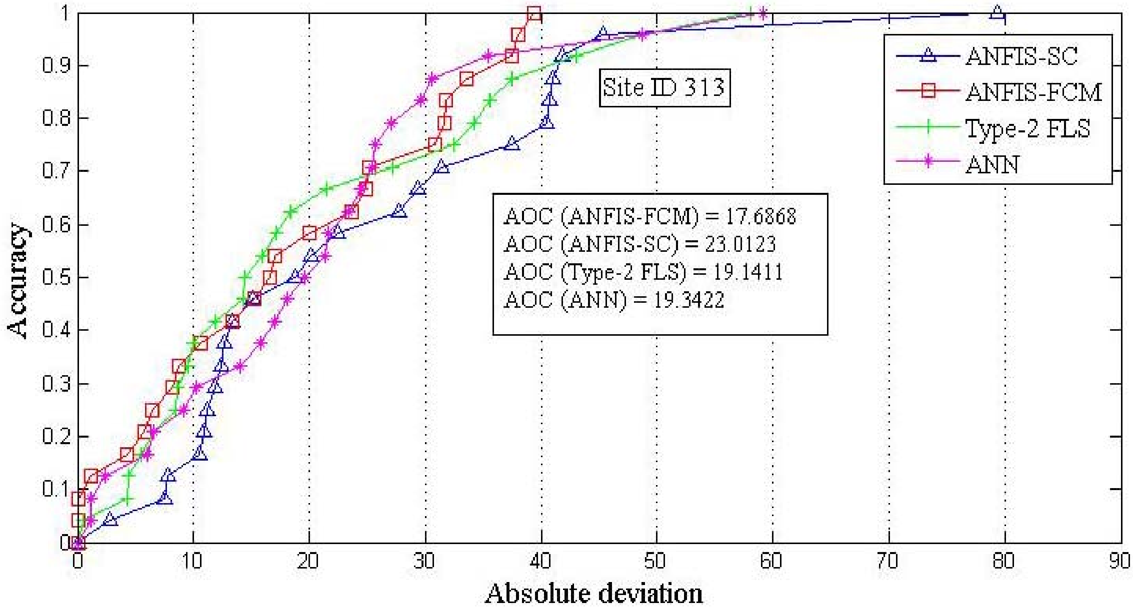


Figure 6-10: REC analysis of different models of the freeway traffic flow prediction at the Station 313.

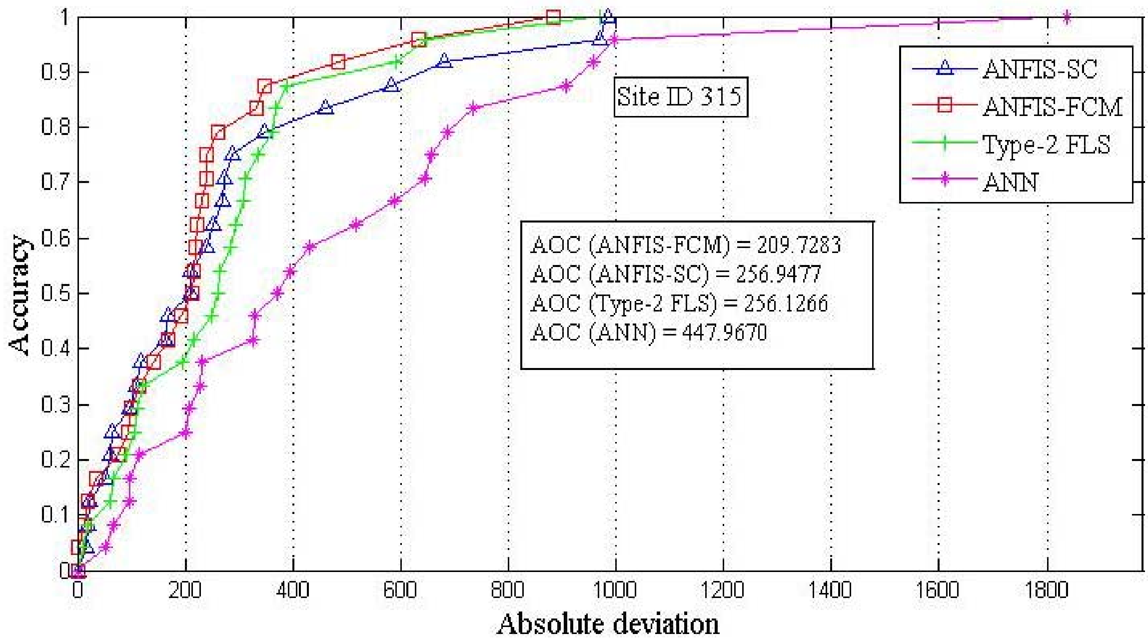


Figure 6-11: REC analysis of different models of the freeway traffic flow prediction at the Station 315.

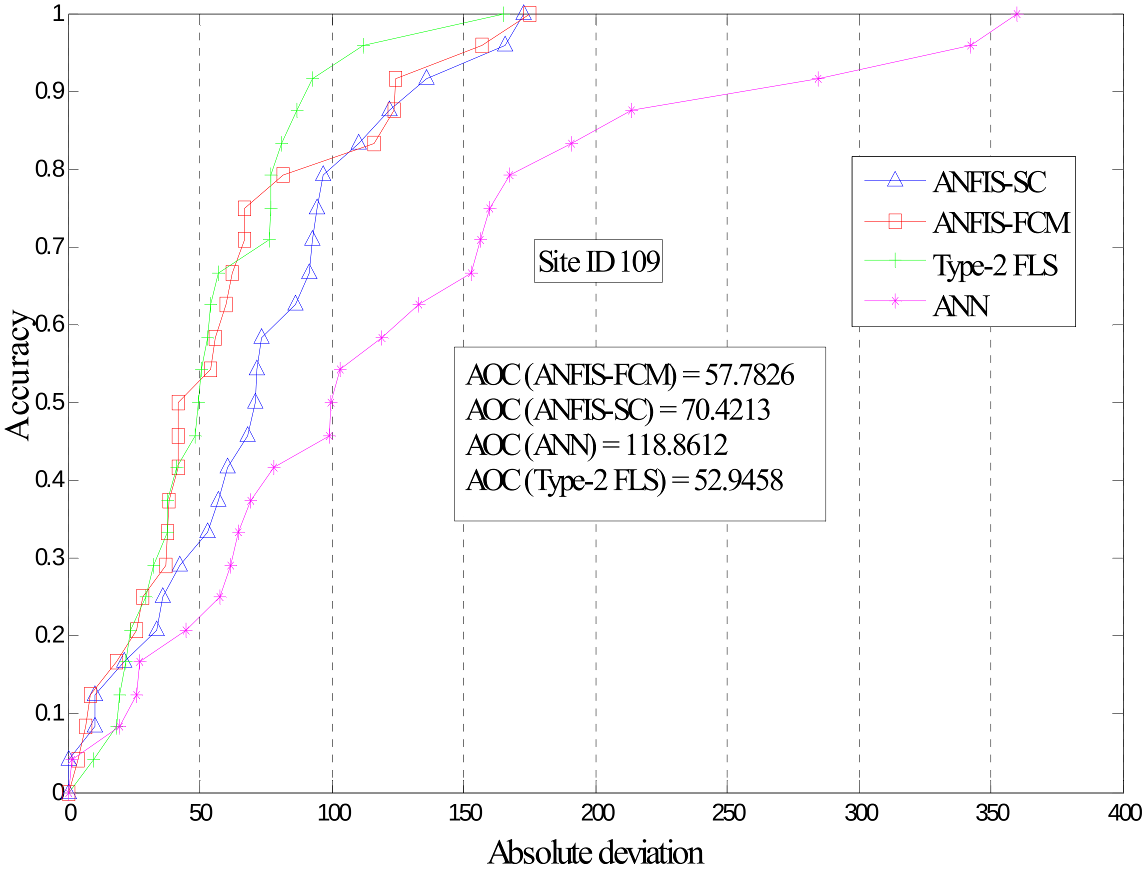


Figure 6-12: REC analysis of different models of the freeway traffic flow prediction at the Station 109.

6.6 Conclusion

This chapter investigated the applicability of advanced AI based models such as ANFIS and type-2 FLS in predicting freeway traffic flow for the local condition of Saudi Arabia. A hybrid model consists of AIM, ANFIS-FCM, ANFIS-SC, type-2 FLS and ANN is proposed in this study. The MAPE of the best models for four freeway sites vary between 5 and 12. The obtained result indicates that the AIM based ANFIS-FCM and type-2 FLS are the valid alternatives for predicting freeway traffic flows in Saudi Arabia. A prediction model based strategy to select freeway sites for future traffic count is also proposed which will benefit MOT in saving cost and improving the understanding the characteristics of the regional traffic. This study introduces REC analysis of all the models in order to provide insights about the performance which is not reported for traffic flow prediction models in the literature.

6.7 Summary

In this chapter a hybrid AI model is proposed to predict freeway traffic flow in four stations. The AIM model selects the input and ANFIS-FCM, ANFIS-SC and type-2 FL models are used to predict the traffic flow. Then the best performing model is selected for prediction. This chapter also provides a strategy for the MOT in order to select future traffic count locations. The performance of the model is evaluated through REC analysis along with traditional error measures such as MAPE, RMSE, MAE etc.

CHAPTER 7 CONCLUSION AND FUTURE RESEARCH

7.1 Conclusion

The recent trend of research focuses on improving real-time adaptive signal control systems, but the wide-scale implementation of such systems seem to be years away particularly in developing countries. On the other hand, the pre-timed traffic signal control does not rely on the sensor data. The efficient operation of pre-timed and actuated traffic signal control depends on the determination of TOD breakpoints. The determination of efficient TOD breakpoints heavily relies on reliable prediction of intersection traffic flow. This research exploited AI based models for traffic flow prediction which are not extensively investigated in the literature such as GMDH based AIM model, type-2 fuzzy logic model and ANFIS models. This research also proposed a novel approach to determine TOD breakpoints.

Type-2 FLS has been successfully applied in wide range of areas and in this study, it is also proven useful for intersection traffic flow prediction. The developed type-2 FL model performs better than traditional singleton and non-singleton type-1 FL model for all approaches with respect to MAE, RMSE and R-square values. The type-2 FL model also ensures better accuracy compared to the FCM and SC based ANFIS model, AIM model and ANN model for the Approach 7 and 8 of the considered intersection “A”. It is also found out that the adoption of SVD reduces the number of rules and complexity of type-2 FL model. This rule reduction technique in type-2 FLS has not been used in traffic flow modeling and prediction problem before. It also contributed in reducing the processing time required for training the model. If the uncertainty associated with the input data is higher, then type-2 FLS model should be used because it is inherently capable to better handle this kind of data.

In order to exploit the potentiality of both ANN and FLS, this study developed ANFIS models to predict intersection traffic flow. For the first time in the literature, the FCM based ANFIS model is investigated for predicting the traffic flow. The ANFIS-FCM performs better than SC based ANFIS model and traditional feedforward NN model. This model even performs better than type-2 model and AIM model in predicting the traffic flow of the Approach 5 and 6. This model can be a good alternative to the traditional ANN and fuzzy logic model in solving similar kind of traffic flow modeling and prediction problems at intersections.

A novel approach of modeling and predicting traffic flow is implemented using GMDH algorithm based AIM model. The developed model outperforms the traditional NN model in terms of MAE, RMSE and R-square values. Due to the self-organizing nature of the model and minimum required interventions, the proposed model is recommended for the local transportation practitioners who are gradually getting exposed to the field of AI based prediction model. This model is deterministic and it also provides mathematical expression.

The developed type-2 FL, ANFIS and abductive network models of this study have specific merit for predicting intersection traffic flow. When the ease of model building is the main concern along with the accuracy then abductive network model is the best choice. If the accuracy of the model is the sole concern, then ANFIS-FCM and type-2 FLS are the favorable models. At last, the type-2 FLS is the proper choice when there is uncertainty associated with the input dataset.

This study proposes a novel approach to determine optimum number and location of TOD breakpoints. It is a subtractive algorithm based K-means method in which the initial cluster centers are obtained through subtractive algorithm. The practitioner can rely on this statistical clustering approach instead of the prevalent judgmental approach.

This study investigated the applicability of advanced AI based models such as ANFIS and type-2 FLS in predicting freeway traffic flow for the local condition of Saudi Arabia. A hybrid model consists of AIM, ANFIS-FCM, ANFIS-SC, type-2 FLS and ANN is proposed in this study. The MAPE of the best models for four freeway sites vary between 5 and 12. The obtained result indicates that the AIM based ANFIS-FCM and type-2 FLS are the valid alternatives for predicting freeway traffic flows. A prediction model based strategy to select freeway sites for future traffic count is also proposed which will benefit MOT in saving cost and improving the understanding the characteristics of the regional traffic. This study introduces REC analysis of all the models in order to provide insights about the performance which is not reported for traffic flow prediction models in the literature.

The traffic flow prediction models proposed in this study are developed based on average 15-min traffic flow data over a month for an average day. These models can’t be readily used for predicting traffic flow data during special events such as accident. If there are significant numbers of training traffic flow data available during special events then similar approach may help in building the models. The proposed surface street and freeway traffic flow prediction models can't be used for other similar locations unless the structure of the models are optimized by using the contextual data. If there are only a few training data are available then the proposed models may not work perfectly like any other machine learning model. In that case, the researchers can think of using ensemble modeling which is discussed in the literature review.

There is a significant difference of driving characteristics between the people of Saudi Arabia and USA. This difference can be significant even for TOD problem solving, if the traffic conditions for same signalized surface streets are modeled in microscopic simulation with the local data of Saudi Arabia and USA. For example, the saturation rate can be higher in case of Saudi Arabia compared to USA. The link speed can be higher in the context of Saudi Arabia even for identical geometric configurations of the road having similar traffic laws and regulations.

7.2 Summary

Although the recent trend of research focuses on developing real-time adaptive signal control systems, the wide-scale implementation of such systems seem to be years away particularly in developing countries. The reasons mainly include associated high costs for implementation and maintenance, and dependency on the sensor data supplied by ITS but the pre-timed traffic signal control do not rely on the sensor data for operation. In pre-timed traffic signal control, the signal timings and cycle lengths may vary depending on TOD to reflect changes in traffic volumes and patterns. They can provide fairly efficient operation during peak periods provided that signal timing settings reflect current conditions. Unfortunately, there exists very little research in the area of determining TOD breakpoints to determine appropriate intervals, or to monitor an existing TOD system to ascertain if the conditions have changed sufficiently to require a new set of plans and/or intervals. These circumstances pose the need of improved performances of widely used pre-timed signal control systems. Even the actuated traffic control systems also require different traffic signal settings in order to cope with the changes of demand during different TOD because a single traffic signal setting doesn't perform well for an entire 24 hour period. But the determination of efficient TOD break points will heavily depend on reliable prediction of intersection traffic flow. In the literature, there are many articles on short term traffic flow prediction aiming to solve the problems associated with adaptive traffic signal controllers. But the traffic flow prediction which may be required for the efficient operation of pre-timed or actuated traffic signal controllers are not addressed in the literature. It can be assumed that the AI based models which are applied for short term prediction can also be used for longer term prediction. This research exploited the AI based models which are tried in short-term traffic flow prediction in few cases or not even attempted.

The evolution of traffic flow in time and space considering complex settings such as signalized streets with closely spaced signals and uncontrolled mid-block demand is not clearly understood to develop accurate mathematical model. Due to the dynamically evolving nature of traffic, the parametric statistical models can model multivariate and multi steps-ahead forecasting relationships but with the cost of reduced accuracy and adaptability, and increased complexity. The limitations of those models inspired the researchers to develop AI based models for traffic flow prediction mainly using ANN. It is capable of learning any smooth nonlinear mapping using numerical data. But it doesn't provide any explicit representation of knowledge about the system. On the other hand, the fuzzy logic based model allows accurate representation of a given system although it can't tackle knowledge stored in the form of numerical data. In order to get benefitted from both the models different methodologies are implemented to develop hybrid neuro-fuzzy model for solving different kinds of prediction problems. In some cases optimization techniques are also adopted to facilitate the hybrid modeling such as evolutionary neuro-fuzzy model, evolutionary support vector regression model. There is also another type of models which are known as committee machines or ensemble models. They refer to the procedures employed to train multiple learning machines and combine their outputs based on the principle that the committee decision, with individual predictions combined appropriately. Many of those proposed models are not investigated in the area of traffic flow prediction specifically for intersection traffic flow prediction. A few examples of those advanced AI models are implemented for short term traffic flow prediction to serve the adopted traffic signal controllers.

This study developed abductive networks of 24-hour traffic flows for an average day of a given month expressed in 15-min intervals. This study exploited the well-proven optimization criteria of AIM in determining network size, element types, connectivity, and coefficients for the optimum model and also investigated the effects of the optionally changeable parameters such as CPM and size of layer 1 on the error measures, aiming to develop better performing networks. The best performing networks were selected, which were developed based on sequential training datasets with different values of CPM, and 6 numbers of limiting layers. The models outperformed the ANN models which were built with the same input datasets. Due to the self-organizing nature of the model and minimum required interventions, the proposed model can be easily used by the practitioners to reduce cost and time in traffic volume collection efforts. This is extremely useful for rapidly developing countries like Saudi Arabia where comprehensive and continuous traffic data collection is still in its infancy. The availability of the mathematical expression of the model will also enhance the acceptability of the model in larger communities.

This study investigated FCM based ANFIS MODEL to predict intersection traffic flow and compared the performance of ANFIS-FCM with ANFIS-SC and NN models. FCM based ANFIS models are not explored in the literature in the area of intersection traffic flow forecasts as per the knowledge of the author. The obtained empirical results by ANFIS-FCM yield more accurate intersection traffic flow forecasting than all the considered models in this study for some specific approaches of an intersection.

This study investigated type-2 FLS for predicting intersection traffic flow which is being explored in wide range of areas. The type-2 fuzzy logic system (FLS) has more design degrees of freedom associated with a type-2 FLS compared to a type-1 FLS which indicates higher potential of better performance. This study proposed type-2 fuzzy models to predict intersection traffic flow and compared the performance of type-2 fuzzy models with type-1 singleton and non-singleton models. The adoption of SVD as a tool to reduce the number of rules and complexity of the model seems promising. It also contributed in reducing the processing time required for training the model.

The performance study of the traffic flow prediction models of the intersection “A”, reveals that the ANFIS-FCM model for eastbound and westbound approaches of the considered intersection outperforms other considered models.

It is observed that for all the models, MAE and RMSE are higher for the Approach 7 and 8 due to the high traffic flow compared to other approaches. The best models are used to determine TOD breakpoints through K-means algorithm. The obtained TOD breakpoints are free from frequent transitions. The proposed approach is systematic in nature and it does not require any subjective intervention.

In order to optimum number of TOD breakpoints a subtractive algorithm based K-means method is used in this study. In this approach the initial cluster centers are obtained through subtractive algorithm. The obtained breakpoints for the observed and predicted traffic flow are evaluated by using SimTraffic. The values of MOEs are compared to investigate the effectiveness and efficiency of using predicted data rather than the original observed data. The use of predicted traffic counts causes increase of total delay (hr), total stops and fuel used (gal) by 3.4, 4.4 and 0.7 percent compared to the corresponding values obtained by using observed traffic counts. The results indicate the potentiality of the use of the predicted data for TOD breakpoint determination and other similar kind of problems.

In this study, a hybrid AI model is proposed to predict freeway traffic flow using local data. The AIM model selects the input and ANFIS-FCM, ANFIS-SC and type-2 FL models are used to predict the traffic flow. Then the best performing model is selected for prediction. This study also provides a strategy for the MOT in order to select future traffic count locations. The performance of the model is evaluated through REC analysis along with traditional error measures such as MAPE, RMSE, MAE etc.

In a nutshell, this study introduces few advanced AI models in modeling and predicting traffic flow for surface streets using foreign data which are not explored or explored in only a very few limited cases. It proposes a clustering based novel approach to determine TOD breakpoints. Then this study proposes a hybrid model for predicting freeway traffic flow based on local data. It also introduces REC analysis to realize the performance of the models along other traditional error measures.

7.3 Future Research

In this section the future research in the areas of traffic flow prediction and TOD breakpoint determination is reported focusing on the investigated models and approaches. The current research in the developed countries is focusing on adaptive traffic controllers. Therefore, the research in developing countries can get benefitted from the current research on adaptive controllers and implement those advanced traffic flow modeling and prediction models in the area of pre-timed and actuated traffic controllers. The same idea can be extended for a hybrid of network having different kind of traffic controllers and traffic management systems.

The GMDH based abductive model can be explored for long term traffic parameter prediction and estimation. The in-built input selection procedure of this model indicates the potential use of this model as input selection for the large and complex road network. Specifically, the researchers in the developing countries can emphasize on this model due to the self-organizing and deterministic nature of this model.

In the literature, there are only a few applications of Neuro-Fuzzy models specifically ANFIS models for traffic flow modeling and prediction problems. As this approach exploit the capabilities of neural network and fuzzy logic systems, the researchers can investigate the performance of FCM based ANFIS models in predicting traffic parameters.

Type-2 fuzzy logic system still didn't get much attention from the researchers of the transportation engineering. But this model is intended to deal with the different sort of uncertainties related to the input data which seem to be very useful for the research in the area of traffic flow modeling and prediction problems.

When individual model fails to perform well for any prediction problem then the ensemble model may perform better than the individual model. In the same line of thought, the investigated models can for ensemble of models to predict traffic flow for critical situations having many missing data and small amount of datasets. There are many available techniques for selecting ensemble models which are not used in the area of traffic flow prediction problems.

In the area of AI, a significant amount of research is emphasizing in building hybrid models by combining different models for different purpose of the model building process. The transportation researchers can emphasize on those advanced models to solve complex traffic flow prediction problems of big network.

The TOD breakpoint determination is still focusing only on small network. The future research may investigate the selection of TOD breakpoint for big and complex network considering traffic parameters such as density, occupancy, speed, and cycle length. Moreover, the future approach should also consider the recurrent congestion condition which is common in urban areas. The investigation of the effects of the congestion on TOD breakpoint selection can be an important contribution in this field of research.

This research paves the path to investigate advanced AI models in predicting freeway traffic flow data for the local condition of the Kingdom of Saudi Arabia. The inspiring results should attract the local researcher to investigate the AI based models in modeling and predicting traffic flow and other traffic parameters using local data.

REFERENCES

[1] Abbas, M. Chaudhary, N. A. Sharma, A. Venglar, S. P. and Engelbrecht, R. J. (2004). "Methodology for determination of optimal traffic responsive plan selection control parameters." *Report, Texas Transportation Institute,* USA.

[2] Abbas, M. (2006). "Optimization of time of day plan scheduling using a multi-objective Evolutionary algorithm." *Proc. Joint International Conference on Computing and Decision Making in Civil and Building Engineering*. Montréal, Canada.

[3] Abdulhai, B. Porwal, H. and Recker, W. (1999). "Short Term Freeway Traffic Flow Prediction Using Genetically-Optimized Time-Delay-Based Neural Networks." *California Partners for Advanced Transit and Highways (PATH). Working Papers: Paper UCB-ITS-PWP-99-1*. URL: http://repositories.cdlib.org/its/path/papers/UCB-ITS-PWP-99-1

[4] Abdulhai, B. Porwal, H. and Recker, W. (1999). "Short Term Freeway Traffic Flow Prediction Using Genetically-Optimized Time-Delay-Based Neural Networks." *California Partners for Advanced Transit and Highways (PATH).* Working Papers: Paper UCB-ITS-PWP-99-1.

[5] Abonyi, J. Feil, B. Nemeth, S. and Arva, P. (2005). "Modified Gath-Geva clustering for fuzzy segmentation of multivariate time-series." *Fuzzy Sets and Systems,* 149(1): 39-56

[6] Abraham, A. (2002). "EvoNF: A Framework for Optimization of Fuzzy Inference Systems Using Neural Network Learning and Evolutionary

Computation." *Proceedings of the 2002 IEEE International Symposium on Intelligent Control,* pp. 327- 332.

[7] Afandizadeh, S. and Kianfar, J. (2009). "A hybrid neuro-genetic approach to short-term traffic volume prediction." *International Journal of Civil Engineering,* 7(1).

[8] Ahmed, S. A., and Cook, A. R. (1979). "Analysis of freeway traffic time-series data using box-Jenkin techniques." *Transportation Research Record 722*, *Transportation Research Board,* Washington, D.C., pp.1–9.

[9] Al Sayyad, M. (1956). "Transportation in Arab countries." Cairo (In Arabic).

[10] Aldagheiri, M. (2009). "The role of the transport road network in the economic development of Saudi Arabia." *Urban Transport, WIT Transactions on the Built Environment*, 107: 275-285.

[11] Andrews, H. C. and Patterson, C. L. (1976). "Singular value decomposition and digital image processing." *IEEE Trans. Acoust., Speech, Signal Processing,* 24:26-53.

[12] Ang, K. K. and Quek, C. (2005). "RSPOP: Rough set-based pseudo-outer-product fuzzy rule identification algorithm." *Neural Comput.,* 17(1):205-243.

[13] Ashok, K. and Ben-Akiva M. (1993). "Dynamic OD matrix estimation and prediction for real-time traffic management systems." *Transportation and Traffic Theory*, Daganzo, C. F. (ed.), Elsevier Science, pp.465-484.

[14] Attaran, M. (1992). "*MSIS - Management Science Information Systems*." New York. John Wiley & Sons.

[15] Babuska, R. (1998). "Fuzzy Modeling for Control." *Kluwer Academic Publishers.*

[16] Bakshi, B. R. (1998). Multiscale PCA with Application to Multivariate Statistical Process Monitoring. *AIChE Journal*, 44(7):1596-1610.

[17] Balkatr, G. (2002). "Evolution-based design of neural fuzzy networks using self-adapting genetic parameters." *IEEE Transactions on Fuzzy Systems,* 10(2):211-221.

[18] Bao-ping, C. and Zeng-qiang, M. A. (2009). "Short-term traffic flow prediction based on ANFIS." *2009 International conference on communication software and networks,* pp. 791-793.

[19] Baranyi, P. and Yam, Y. (1997) "Singular value-based approximation with non-singleton fuzzy rule base." *In Proc. 7th Int. Fuzzy Systems Association World Congress. IFSA'97,* Prague, pp. 127-132.

[20] Baranyi, P. Yam, Y. Varkonyi-Koczy, A. R. Patton, R. J. (2003). "SVD-based reduction to MISO TS models." *IEEE transactions on industrial electronics,* 50(1): 232-242.

[21] Beaver, S. and Palazoglu, A. (2007). "Cluster analysis for auto-correlated and cyclic chemical process data." *Industrial & Engineering Chemistry Research,* 46(11):3610-3622.

[22] Ben-Akiva, M. Cascetta, E. and Whittaker, J. (1993). "Recent progress in short-range traffic prediction." *Compendium of Technical Papers, ITE,* pp. 262-265.

[23] Bi, J. and Bennett, K. P. (2003). "Regression error characteristic curve." P*roc. Twentieth Int. Conf. Machine Learning (ICML'03).*

[24] Boser, B. E. Guyon, I. M. and Vapnik, V. N. (1992). "A training algorithm for optimal margin classifiers." *In D. Haussler, editor, Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory,* pp. 144-152, Pittsburgh, PA, ACM Press.

[25] Box, G. E. P. and Jenkins, G. M. (1976). "Time series analysis: forecasting and control." 2nd edition, Holden-Day, San Francisco.

[26] Boxill, S. A. and Yu, L. (2000) "An Evaluation of Traffic Simulation Models for Supporting ITS Development," *Center for Transportation Training and Research,* Texas Southern University, USA.

[27] Braun, H. and Weisbrod, J. (1991). "Evolving neural feedforward networks." *International conference on Artificial Neural Nets and Genetic Algorithms (ANNGA93),* Innsbruck, Austria, pp. 13-22.

[28] Brown, G. Ensemble Learning. (2010). *"Encyclopedia of Machine Learning,"* C.Sammut & G.I.Webb (Eds.), Springer Press.

[29] Brown, G. Wyatt, J. L. Harris, R. and Yao, X. (2005). "Diversity Creation Methods: A Survey and Categorization." *Journal of Information Fusion, 6(1):5-20.*

[30] Burges, C. J. C. and Scholkopf, B. (1997). "Improving the accuracy and speed of support vector learning machines." *In Mozer, M. Jordan, M. and Petsche, T. editors, Advances in Neural Information Processing Systems 9,* pp. 375-381, Cambridge, MA, MIT Press.

[31] Burghout, W. (2004). "Hybrid Microscopic-Mesoscopic Traffic Simulation," *Doctoral Dissertation,* Royal Institute of Technology, Stockholm, Sweden.

[32] Caruana, R. (1997). "Multitask learning." *Machine Learning,* 28(1), 41-75.

[33] Castro-Neto, M. Jeong, Y.-S. Jeong, M.-K., Han, L. D. (2009). "Online-SVR for Short-Term Traffic Flow Prediction under Typical and Atypical Traffic Conditions." *Expert Systems with Applications*, 36:6164-6173.

[34] Chaturvedi, D. K. (1997). "Modeling and simulation of electrical power systems: an alternative approach." *Doctoral Thesis,* Dayalbagh Educational Institute, Dayalbagh, Agra, India.

[35] Chaturvedi, D. K. (2008). "Soft computing techniques and its applications in electrical engineering." Kacprzyk, J. (ed.) *Studies in Computational Intelligence*, vol. 103, Springer.

[36] Chen, H. and Grant-Muller, S. (2001). "Use of sequential learning for short-term traffic flow forecasting." *Transportation Research Part C*, 9:319-336.

[37] Chen, S. Wang, W. Ren, G. (2005). "A hybrid approach of traffic volume forecasting based on wavelet transform, neural network and Markov model." *IEEE International Conference on Systems, Man and Cybernetics***,** 1:393- 398.

[38] Chen, S. Wang, W. Ren, G. (2005). "A hybrid approach of traffic volume forecasting based on wavelet transform, neural network and Markov model." *IEEE International Conference on Systems, Man and Cybernetics*, 1:393- 398.

[39] Chen, Y. H. Yang, B. and Dong, J. W. (2006). "Time-series prediction using a local linear wavelet neural network." *Neurocomputing,* 69:449-465.

[40] Chen, Y. Zhang, Y. Hu, J. and Li, X. (2006). "Traffic data analysis using kernel PCA and self-organizing map." *Intelligent Vehicles Symposium 2006,* June 13-15, Tokyo, Japan.

[41] Cheng, M.-Y. and Ko, C. H. (2006). "A genetic-fuzzy-neuro model encodes FNNs using SWRM and BRM." *Engineering Applications of Artificial Intelligence,* 19(8):891-903.

[42] Cheslow, M., Hatcher, S. G., and Patel, V. M. (1992). "An initial evaluation of alternative intelligent vehicle highway systems architecture." *Rep. No. 92w0000063,* MITRE Corporation, Bedford, Mass.

[43] Chiang, M. M-T., and Mirkin, B. (2007). "Experiments for the number of clusters in K-means." *Proc. 13th Portuguese Conference on Progress in Artificial Intelligence,* Portugal, pp. 395-406.

[44] Chiu, S. (1994). "Fuzzy Model Identification Based on Cluster Estimation." *Journal of Intelligent & Fuzzy Systems,* 2(3).

[45] Comon, P. and Golub, G. H. (1990). "Tracking a few extreme singular values and vectors in signal processing." *Proc. IEEE 1990,* 78:1327-1343.

[46] Cordón, O. Gomide, F. Herrera, F. Hoffmann, F. and Magdalena, L. (2004). "Ten years of genetic fuzzy systems: current framework and new trends." *Fuzzy Sets and Systems,* 141(1):5-31.

[47] Cui, W. Z. Zhu, C. C. and Zhao, H. P. (2005). "Prediction of thin film thickness of field emission using wavelet neural networks." *Thin Solid Films,* 473:224-229.

[48] Danech-Pajouh, M. and Aron, M. (1991). "ATHENA, a method for short-term inter-urban traffic forecasting." *Report 177, INRETS,* Arcueil, Paris.

[49] Dash, P. K. Mishra, S. Dash, S. and Liew, A. C. (2000). "Genetic optimization of a self organizing fuzzy-neural network for load forecasting." *Proceedings of the IEEE Power Engineering Society Winter Meeting,* 2:1011-1016.

[50] Davies, D.L., Bouldin, D.W. (1979). "A cluster separation measure." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1(4): 224-227.

[51] Davis, G. A. (1990). "Adaptive forecasting of freeway traffic congestion." *Transportation Research Record,* No.1287, pp. 29-33.

[52] Davis, G. A., and Nihan, N. L. (1991). "Nonparametric regression and short term freeway traffic forecasting." *J. Transp. Eng.,* 117(2): 178–188.

[53] Der Voort, M. Dougherty, M. and Watson, S. (1996). "Combining Kohonen maps with ARIMA time series models to forecast traffic flow." *Transportation Research Part C,* 4(5):307-318.

[54] Dia, H. (2001). "An object-oriented neural network approach to short-term traffic forecasting." *European Journal of Operation Research,* 131(2):253-261.

[55] Ding, A. Zhao, X. Jiao, L. (2007). "Traffic flow time series prediction based on statistics learning theory." *The IEEE 5th International Conference on Intelligent Transportation Systems,* 20(6):95-99.

[56] Disbro, J. E. Frame, M. (1989). "Traffic flow theory and chaotic behavior." *Transportation Research Record,* 8(1):109-115.

[57] Dongli, W. Yan, Z. Xiaoyang, H. (2007). "Radial Basis Function Neural Network-Based Model Predictive Control for Freeway Traffic System."

International Journal of Intelligent Systems Technologies and Applications. 2(4):370-388.

[58] Dubois, D. and Prade, H. (1982). *"Fuzzy Sets and Systems: Theory and Applications."* New York, Academic.

[59] Dunn, J.C. (1974). "Well separated clusters and optimal fuzzy partitions." *Cybernetics and Systems,* 4(1):95-104.

[60] Durbin, J. (2000). "The Foreman lecture: the state space approach to time series analysis and its potential for official statistics." *Australian and New Zealand Journal of Statistics,* 42 (1), 1–23.

[61] Eberhart, R. C. (1992). "The role of Genetic Algorithms in Neural Network Query-based Learning and Explanation Facilities." *IEEE International Workshop on Combinations of Genetic Algorithms and Neural Networks (COGANN-92),* Baltimore, pp. 169-183.

[62] Fakhreddine, K. and de Silva, C. (2004). "Soft computing and tools of intelligent systems design." Pearson, UK.

[63] Farag, W. A. Quintana, V. H. and Lambert-Torres, G. (1998). "A genetic-based neuro-fuzzy approach for modeling and control of dynamical systems**."** *IEEE Transactions on Neural Networks,* 9(5):756 – 767.

[64] Fawcett, T. (2003). "ROC graphs: notes and practical considerations for data mining." *Technical report, HP-2003-4.*

[65] Federal Highway Administration. (2001). "Traffic monitoring guide." URL: http://www.fhwa.dot.gov/ohim/tmguide/

[66] Fellendorf, M. (1994). “VISSIM: A microscopic simulation tool to evaluate actuated signal control including bus priority,” *64th ITE Annual Meeting.*

[67] Ferguson, E. Travel demand management and public policy, Ashgate, Aldershot, 2000.

[68] Friesz, T. L. and Bernstein, D. (200). “Analytical dynamic traffic assignment models.” *Handbook of Transport Modeling,* Hensher, D. A. Button, K. J. (ed.), Pergamon, Amsterdam, pp. 181-195.

[69] Fukunaga, K. and Koontz, W. (1970). “Application of Karhunen-Loeve expansion to feature selection and ordering.” *IEEE Trans. Comput., C-19, 311.*

[70] Gartner, N. H. Stamatiadis, C. and Tarnoff, P. J. (1995). Development of advanced traffic signal control strategies for intelligent transportation systems: multilevel design. *Transportation Research Record,* 1494:98–105.

[71] Gartner, N. H. Tarnoff, P. J. and Andrews, C. M. (1991). “Evaluation of Optimized Policies for Adaptive Control Strategy.” *Transportation Research Record 1324,* TRB, National Research Council, Washington, D. C.

[72] Giuliano, G. and Wachs, M. (1992). “Responding to congestion and traffic growth: transportation demand management.” The University of California, Transportation Center.

[73] Goodman, L., Kruskal, W. (1954). “Measures of associations for cross-validations.” *Journal of the American Statistical Association,* 49:732-764.

[74] Goonatilake, S. and Sukhdev, S. (1995). “Intelligent Hybrid Systems.” S. Goonatilake and S. Khebbal (ed.), Wiley Publishing, Toronto.

[75] Gordon, R. L. Reiss, R. A. Haenel, H. Case, E. R. French, R. L. Mohaddes, A. and Wolcutt, R. (1996). "Traffic Control Systems Handbook." Publication Number FHWA-SA-95-032, U. S. Department of Transportation, USA.

[76] Gui-Yan, J. (2004). "Technologies and application of the identification of road traffic condition." *Beijing: Communications Press,* (in Chinese).

[77] Guo, J. (2005). "Adaptive estimation and prediction of univariate vehicular traffic conditions series." *Ph. D. Dissertation,* North Carolina State University.

[78] Hai-shuang, G. Wen-ge, M. A. and Yan-yu, M. (2008). "Traffic flow prediction based on hierarchical genetic optimized algorithm." *The 3rd International Conference on Innovative Computing Information and Control (ICICIC'08). IEEE.*

[79] Halgamuge, S. K. and Glesner, M. (1994). "Neural networks in designing fuzzy systems for real world applications." *Int. J. Fuzzy Sets Syst.*, 65(1):1-12.

[80] Hamed, M. M. Al-Masaeid, H. R. and Said, Z. M. B. (1995). "Short-Term Prediction of Traffic Volume in Urban Arterials." *J. Transp. Engrg. 121(3):249-254.*

[81] Hammarling, S. J. (1985). "The singular value decomposition in multivariate statistics." *ACM Signum Letter*, 20:l2-25.

[82] Han, C. and Song, S. A review of some main models for traffic flow forecasting, IEEE, 2003, pp. 216-219.

[83] Hanks, J. and Lomax, T. (1991). "Roadway congestion in major urban areas: 1982 to 1988." *Paper presented at the 70th Annual Meeting,* Transportation Research Board, Washington, D.C.

[84] Hansen, J. V. McDonald, J. B. and Nelson, R. D. (2002). "Time series prediction with genetic-algorithm designed neural networks: an empirical comparison with modern statistical models." *Computational Intelligence, 15(3):171-184.*

[85] Hassoun, M. H. (1995). "Fundamentals of Artificial Neural Networks," *MIT Press.*

[86] Hauser, T., Scherer, W. T., and Smith, B. L. (2000). "Signal system data mining." *Report, Center for Transportation Studies,* University of Virginia, USA.

[87] Hayashi, I., Nomura, H., Yamasaki, H. and Wakami, N. (1992). Construction of fuzzy inference rules by NDF and NDFL. *Intl. J. Approx. Reason.,* 6(2):241-266.

[88] Haykin, S. (1994). "Neural Networks: A comprehensive foundation." Prentice-Hall, Upper Saddle River, N. J.

[89] He, S., Hu, C., Song, G. -J., Xie, K. -Q., and Sun, Y. -Z. (2008). "Real-Time Short-Term Traffic Flow Forecasting Based on Process Neural Network." *Advances in Neural Networks,* 5264:560-569, Springer Berlin/Heidelberg.

[90] Head, K. L.(1995). "An Event-Based Short-Term Traffic Flow Prediction Model." *Transportation Research Record 1510,* pp.45-52.

[91] Hellinga, B. R. (1997). "Estimating dynamic origin-destination demands from link and probe counts." *Transportation Research A,* 31(1), 83.

[92] Hellmann, M. (2001). "Fuzzy Logic Introduction." Université de Rennes, France.

[93] Herrera, F. (2008). “Genetic fuzzy systems: taxonomy, current research trends and prospects.” *Evol. Intel.* 1:27–46

[94] Herrera, F. and Verdegay, J. L. (1996). (ed.) Genetic Algorithms and Soft Computing. *Number 8 in Studies in Fuzziness and Soft Computing.* Physica-Verlag.

[95] Hibbs, R. A. (1994). “Speeding up backpropagation: a comparative study.” Technical Report, Knowledge-based Engineering Systems Groups, University of South Australia, Australia.

[96] Holland, J. H. (1975). “Adaption in natural and aritifical systems.” University of Michigan Press, Ann Arbor, 1975.

[97] Holland, J. H. and Reitman, J. S. (1978). “Cognitive systems based on adaptive algorithms.” *In D. A. Waterman and F. Hayes-Roth, editors,* Pattern-directed inference systems. Academic Press.

[98] Hongqiang, L. I. Huapu, L.U. and Weihua, Z. (2005). “A new optimization method for time-of-day signal timing transition of arterial traffic.” *Proc. of the Eastern Asia Society for Transportation Studies,* 5:1352 – 1357.

[99] Hoogendoorn, S. Hoogendoorn-Lanser, S. Schuurman, H. (1999). “Fuzzy perspectives in traffic engineering.” Workshop on Intelligent Traffic Management Models, Delft.

[100] Hornik, K. M. Stinchcombe, M. and White, H. (1989). “Multilayer feed-forward networks are universal approximators.” *Neural Networks,* 2(5):.359-366.

[101] Hu, C. Xie, K. Song, G. and Wu, T. (2008). “Hybrid process neural network based on spatio-temporal similarities for short-term traffic flow prediction.

Proceedings of the 11th International IEEE Conference on Intelligent Transportation Systems, Beijing, China, October 12-15, pp. 253-258

[102] Hu, C. Xie, K. Song, G. and Wu, T. Hybrid Process Neural Network based on Spatio-Temporal Similarities for Short-Term Traffic Flow Prediction. *The 11th International IEEE Conference on Intelligent Transportation Systems, 2008.*

[103] Hua, J. and Faghri, A. (1993) "Dynamic traffic pattern classification using artificial neural networks." *Transportation Research Record: Journal of the Transportation Research Board,* No. 1399, pp. 14-19.

[104] Huang, S. and Sadek, A. W. (2009). "A novel forecasting approach inspired by human memory: the example of short-term traffic volume forecasting." *Transportation Research Part C,* 17:510-525.

[105] Huang, Y. (2008). "Short-term traffic flow forecasting based on Wavelet network model combined with PSO." *Proceedings of the 2008 International Conference on Intelligent Computation Technology and Automation,* pp. 249-253.

[106] Huang, Y. Gedeon, T. D. and Wong, P. M. (2001). "An integrated neural-fuzzy-genetic-algorithm using hyper-surface membership functions to predict permeability in petroleum reservoirs." *Engineering Applications of Artificial Intelligence,* 14(1):15-21.

[107] Hubert, L. and Schultz. J. (1976). "Quadratic assignment as a general data-analysis strategy." *British Journal of Mathematical and Statistical Psychology,* 29:190-241.

[108] Hunt, P. B. Robertson, D. I. Bretherton, R. D. and Winton, R. I. (1981). "SCOOT- A traffic responsive method of coordinating signals." *Report LR 1014.* Transport and Road Research Laboratory, Crowthorne, Berkshire, England.

[109] Ichimura, T. Tazaki, E. (1995). "Applying adaptive structured genetic algorithm to reasoning and learning method for fuzzy rules using neural networks," *Neural Networks, Proceedings of IEEE International Conference on,* 6:3124-3128.

[110] Ishak, S. and Trifirò, F. (2007). "Neural Networks." *Transportation Research Circular, Artificial Intelligence in Transportation*. Number E-C113, pp. 17-33.

[111] Ishibuchi, H. Tanaka, H. and Okada, H. (1994). "Interpolation of fuzzy if-then rules by neural networks." *Int. J. Approx. Reason.,* 10(1):3-27.

[112] Ishigami, H. Fukuda, T. Shibata, T. and Arai, F. (1995). "Structure optimization of fuzzy neural network by genetic algorithm." *Fuzzy Sets and Systems,* 71(3): 257-264.

[113] Ivanciuc, O. (2007). "Applications of Support Vector Machines in Chemistry."*In Reviews in Computational Chemistry,* vol. 23, Eds: K. B. Lipkowitz and T. R. Cundari. Wiley-VCH, Weinheim, pp. 291-400.

[114] Jaccard. P. (1912). "The distribution of flora in the alpine zone." *New Phytologist,* vol. 11, pp. 37–50.

[115] Jacob, C. and Rehder, J. (1993). "Evolution of neural net architectures by a hierarchical grammar-based genetic system." *International conference on*

artificial neural nets and genetic algorithms (ANNGA93), Innsbruck, Austia, pp. 72-79.

[116] Jang, J.-S. R. (1993). “ANFIS: Adaptive-Network-Based Fuzzy Inference System.” *IEEE Transactions on Systems, Man, and Cybernetics,* 23(3):665-685.

[117] Jang, J.-S.R. (1996). “Input selection for ANFIS learning.” *In Proceedings of the IEEE International Conference on Fuzzy Systems,* New Orleans, 1996.

[118] Janten, J. (1998). “Neuro-Fuzzy Modeling.” *Technical University of Denmark,* Department of Automation, Denmark. Tech. report no 98-H-874.

[119] Janten, J. Neuro-Fuzzy Modeling. *Technical University of Denmark, Department of Automation, Denmark. Tech. Report no 98-H-874, 30 Oct 1998.*

[120] Jiang, X. and Adeli, H. (2005). “Dynamic wavelet neural network model for traffic flow forecasting.” *Journal of Transportation Engineering,* 131:771-779.

[121] Jianyu, Z. Lei, J. Yuchui, C. and Yong, Z. (2007). “Traffic Flow Forecasting Model of Signalized Intersections Based on Particle Swarm Optimization.” *Control Conference,* Chinese, pp. 38-40.

[122] Jiawei, H. and Micheline, K. (2000). “Data Mining: Concepts and Techniques.” Academic Press.

[123] Jin, F. and Sun, S. (2008). “Neural Network Multitask Learning for Traffic Flow Forecasting.” *International Joint Conference on Neural Networks (IJCNN 2008).*

[124] Joback, K. G. (1984). “A unified approach to physical property estimation using multivariate statistical techniques.” *MS Thesis,* Mass. Inst. of Tech.

[125] Johanns, R. D. and Roozemond, D. A. (1993). An object based traffic control strategy: a chaos theory approach with an object-oriented implementation. *Advanced Technologies,* 4(2): 231-242.

[126] Johansen, T. A. (1995). “On the optimality of the Takagi-Sugeno-Kang fuzzy inference mechanism.” *In Proc. FUZZ-IEEE/IFES’95,* Yokohama, Japan, March 1995, pp. 97-102.

[127] John, R. I. Innocent, P. R. and Barnes, M. R. (1998). “Type 2 fuzzy sets and neuro-fuzzy clustering of radiographic tibia images.” *In Proc. 1998 IEEE Int. Conf. Fuzzy Syst.,* Anchorage, AK, May 1998, pp. 1373-1376.

[128] Ju, Y. J. Kim, C. E. and Shim, J. C. (1997). “Genetic-based fuzzy models: Interest rate forecasting problem.” *Computers & Industrial Engineering.* 33(3-4):561-564.

[129] Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. Transaction of ASME. *Journal of Basic Engineering,* 82:35-45.

[130] Kardanpour, Z. Hemmateenejad, B. and Khayamian, T. (2005). Wavelet neural network-based QSPR for prediction of critical micelle concentration of Gemini surfactants. *Analytica Chimica Acta,* 531: 285-291.

[131] Karmer, M. A. (1991). “Nonlinear principal component analysis using auoassociative neural networks. *AIChE Journal,* 37(2):233-243.

[132] Karnik, N, N. and Mendel, J. M. (1999). “Application of type-2 fuzzy logic systems to forecasting of time series.” *Inform. Sci.,* 120: 89-111.

[133] Karnik, N, N. and Mendel, J. M. (1988). "An introduction to type-2 fuzzy logic systems." Univ. of Southern California, Los Angeles, CA, June 1988.

[134] Karnik, N, N. and Mendel, J. M. (1998). "Introduction to type-2 fuzzy logic systems." In Proc. *IEEE FUZZ Conf., Anchorage,* AK.

[135] Karnik, N, N. and Mendel, J. M. (1998). "Type-2 fuzzy logic systems: type reduction, in *IEEE Syst., Man, Cybern. Conf.,* San Diego, CA.

[136] Karnik, N, N. and Mendel, J. M. (2000). "Operations on type-2 fuzzy sets." *Fuzzy Sets Syst.*

[137] Karnik, N. N. and Mendel, J. M. (1999). "Applications of type-2 fuzzy logic systems: handling the uncertainty associated with surveys." In Proc. *FUZZ-IEEE Conf.,* Seoul, Korea, Aug.

[138] Kaysi, I. and Ben-Akiva, M. and Koutsopoulos, H. (1993). "Integrated approach to vehicle routing and congestion prediction for real-time driver guidance." *Transportation Research Record,* 1408:66-74.

[139] Keller, J. M. Yager, R. R. and Tahani, H. (1992). "Neural network implementation of fuzzy logic. *Fuzzy sets Syst.,* 45(1):1-12.

[140] Kim, D. (1997). "Forecasting time series with a genetic fuzzy predictor ensemble." *IEEE Transactions on Fuzzy Systems,* 5(4):523-535.

[141] Kirby, H. R. Watson, S. M. and Dougherty, M. S. (1997). "Should we use neural networks or statistical models for short term motorway traffic forecasting?" *International Journal of Forecasting,* 13:43-50.

[142] Kivikunnas, S. (1998). "Overview of process trend analysis methods and applications." *ERUDIT workshop on applications in pulp and paper industry,* CD ROM.

[143] Kresta, J. MacGregor, J. F. and Marlin, T. E. (1991). "Multivariate Statistical Monitoring of Process Operating Performance." *Can. J. Chem. Eng.*, 69:35-47.

[144] Kun, H. Enfa, C. S. and Zhenguo, Z. (2003). "Research on a nonlinear chaotic prediction model for urban traffic flow." *Journal of Southeast University,* 19(4): 71-75.

[145] Kuo, R. J. (2001). "A sales forecasting system based on fuzzy neural network with initial weights generated by genetic algorithm." *European Journal of Operational Research,* 129(3): 496-517

[146] Lamrous, S. and Taileb, M. (2006). "Divisive hierarchical K-Means." Proc. *Conference on Intelligent Agents, Web Technologies and Internet Commerce,* pp. 18 – 18

[147] Last, M. Klein, Y. and Kandel, A. (2000). "Knowledge discovery in time-series databases." *IEEE Transactions on Systems, Man, and Cybernetics,* 31(1):160-169.

[148] Laub, A. J. (1985). "Numerical linear algebra aspects of control design computations." *IEEE Trans. Automat. Contr.*, 30: 97-108.

[149] Lebacque, J. (1983). "Simulation semi macroscopique des reseaux urbains." *Technical report, Institut de Recherche sur les Transports.*

[150] Ledoux, C. (1997). “An urban traffic flow model integrating neural networks.” *Transpns. Res.-C.* 5(5):287-300.

[151] Lee, D. -H. Jeng, S. -T. and Chandrasekar, P. (2004a). “Applying Data Mining Techniques for Traffic Incident Analysis.” *Journal of The Institution of Engineers,* Singapore, 44(2):90-102.

[152] Lee, D. H. Zheng, W. Z. and Shi, Q. X. (2004b). Short-term freeway traffic flow prediction using a combined neural network model. *CDROM of the 83rd TRB Annual Meeting,* January 11–15.Washington, D.C.

[153] Lee, M. Lee, S. Y. and Park, C. H. (1994). “A new neuro-fuzzy identification model of nonlinear dynamic systems.” *Int. J. Approx. Reason.,* 10(1):3-27.

[154] Lee, S. and Fambro, D. B. (1999). “Application of subset autoregressive integrated moving average model for short-term freeway traffic volume.” *Journal of the Transportation Research Board,* 1678:179-188.

[155] Lee, S. and Lee, K. H. (2002). “A ranking method for Type-2 fuzzy values". *Journal of Korea Fuzzy and Intelligent Systems Society,* 12(4):341-346.

[156] Levin, M. and Tsao, Y. -D. (1980). “On forecasting freeway occupancies and volumes.” *Transportation Research Record*, 773: 47-49.

[157] Li, L. Lin, W.- H. and Liu, H. (2005). “Type-2 fuzzy logic approach for short-term forecasting.” IEE Proc. *Intell. Transp. Syst.,* 153(1).

[158] Liang, Q. and Mendel, J. M. (2000c). “Interval type-2 fuzzy logic systems: theory and design.” *IEEE Trans Fuzzy Systems,* 8(5):535-550.

[159] Liang, Q. Mendal, J. M. (2001). "MPEG VBR video traffic modeling and classification using fuzzy techniques." *IEEE Trans. on Fuzzy System,* 9(1):183-193.

[160] Liang, Q. Mendel, J. M. (2000b). "Overcoming time-varying co channel interference using type-2 fuzzy adaptive filter." *IEEE Trans. On Circuits and Systems,* 47:1419–1428.

[161] Liang, Q. Mendel, J. M. (2000d). "Designing interval type-2 fuzzy logic systems using SVD-QR method: rule reduction." *International journal of intelligent systems,* 15: 939-957.

[162] Lin, C. T. and Lee, C. S. (1991). "Neural-network-based fuzzy logic control and decision system. *IEEE Trans. Comput.,* 40(12):1320-1336.

[163] Lovmar, L. Ahlford, A. Jonsson, M. and Syvänen, A. -C. (2005). "Silhouette scores for assessment of SNP genotype clusters." *BMC Genomics*, 6(1), Record No. 35.

[164] MacGregor, J. F. (1994). "Statistical Process Control of Multivariate Processes." *IFAC ADCHEM,* Kyoto, Japan.

[165] Mamdani, E. H. and Assilian, S. (1975). "An experiment in linguistic synthesis with a fuzzy logic controller." *International J. Man-Mach. Stud.,* 7:1-13.

[166] Mauro, V. and Di Taranto, D. (1990). "UTOPIA." *Proc. 6th IFAC/IFIP/IFORS Symposium on Control, Computers and Communication in Transportation, Paris, France.*

[167] Ming-bao, P. and Xin-ping, Z. (2008). "Traffic flow prediction of chaos time series by using subtractive clustering for fuzzy neural network modeling."

Second international symposium on Intelligent Technology Application, IEEE, pp. 23-27.

[168] Mirchandani, P and Head, L (2001). A real-time traffic signal control system: architecture, algorithms, and analysis. *Transportation Research Part C: Emerging Technologies,* 9(6):415–32.

[169] Mitra, S. and Hayashi, Y. (2000). "Neuro-fuzzy rule generation: survey in soft computing framework." *IEEE transactions on neural networks,* 11(3):748-768.

[170] MOP. (1975). "The second development plan." *Ministry of Planning*, Saudi Arabia, MOP Press, Riyadh.

[171] MOP. (1980). "The third development plan." *Ministry of Planning*, Saudi Arabia, MOP Press, Riyadh, 1980.

[172] MOP. (1985). "The fourth development plan." *Ministry of Planning*, Saudi Arabia, MOP Press, Riyadh.

[173] MOP. (1990). "The fifth development plan." *Ministry of Planning*, Saudi Arabia, MOP Press, Riyadh.

[174] MOP. (2000). "Statistical yearbook." Central Department of Statistics, *Ministry of Planning*, Riyadh.

[175] MOP. (2000). Ministry of Planning, The seventh development plan, MOP Press, Riyadh, 2000.

[176] MOT. (1971). "Roads and ports in Saudi Arabia." *Ministry of Transportation*, MOT Press, Riyadh, 1971.

[177] MOT. (1985). "The roads in Saudi Arabia Ministry of Transportation." MOT Press, Riyadh, 1985.

[178] MOT. (1986). “Transport and communication progress and achievements.” Ministry of Transportation, Riyadh.

[179] MOT. (2003). “The ways of development, Saudi Arabia.” *Ministry of Transportation*, Riyadh.

[180] MOT. (2006). “Annual traffic data report Ministry of Transportation.” *Road Service Department*, MOT, Saudi Arabia.

[181] MOT. (2007). “Annual traffic data report Ministry of Transportation.” *Road Service Department*, MOT, Saudi Arabia.

[182] Muhammad, A. and King, G.A. (1997). “Foreign exchange market forecasting using evolutionary fuzzy networks. Computational Intelligence for Financial Engineering (CIFEr).” *Proceedings of the IEEE/IAFE 1997,* New York City, NY, USA, pp. 213-219.

[183] Musavi, M. T. Chan, K. H. and Kalantri. K. (1994). “On generalization ability of neural network classifiers.” *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 16(6):659-663.

[184] Mussa, R. and Selekwa, M. (2003). “Proposed methodology of optimizing transitioning between time-of-day timing plans.” *Journal of Transportation Engineering,* 129(4):392-398.

[185] Nauck, D. and Klawonn, F. (1996). “Neuro-Fuzzy Classification Initialized by Fuzzy Clustering.” *Proc. Fourth European Congress Intelligent Techniques and Soft Computing (BUFIT’96),* Aachen.

[186] Nayak, P. C. Sudheer, K. P. Rangan, D. M. and Ramasastri, K. S. (2004). "A neuro-fuzzy computing technique for modeling hydrological time series." *Journal of Hydrology,* 291(1-2):52-66.

[187] Ni, D. (2001). "Panorama of Transportation Simulation," *Georgia Institute of Technology, USA, 2001, URL: http://gte710q.tripod.com/Research/panorama.pdf*

[188] Nicholson, H. and Swann, C. D. (1974). "The prediction of traffic flow volumes based on spectral analysis." *Transpn. Res.* 8:533-538.

[189] Niemeier, D. A. Utts, J. M. and Fay, L. (2002). "Cluster analysis for optimal sampling of traffic count data: air quality example." *J. Transp. Engrg,* 128(1):97-102.

[190] Nomikos, P. and MacGregor, J. F. (1994). "Monitoring batch processes using multiway principal component analysis." *AIChE J., 40(8): 1361-1375.*

[191] Okutani, I. and Stephanedes, Y. J. (1984). "Dynamic prediction of traffic volume through Kalman filtering theory." *Transportation Research, Part B,* 18(1):1-11.

[192] Ortuzar, S. Dios, J. D. and Willumsen, L. G. (2001). "Modeling Transport (3rd Edition)." J. Wiley, Chichester.

[193] Pai, P. -F.; Lin, K. -P. and Chang, P. -T. (2007). "Potential assessment of an ellipsoidal neural fuzzy time series model for freeway traffic prediction." *Adaptive and Natural Computing Algorithms,* 4431: 631-639, Springer Berlin / Heidelberg.

[194] Park, B. and Lee, J. (2008). "A procedure for determining time-of-day breakpoints for coordinated actuated traffic signal systems." *KSCE Journal of Civil Engineering,* 12(1):37-44.

[195] Park, B. B. (2002). "Hybrid neuro-fuzzy application in short-term freeway traffic volume forecasting." *Transportation Research Record: Journal of the Transportation Research Board,* 1802:190-196.

[196] Park, B. Lee, D. and Yun, I. (2003a). "Enhancement of time of day based traffic signal control." *Proc. of the IEEE International Conference on Systems, Man and Cybernetics.* 4:3619-3624.

[197] Park, B. Messer, C. J. and Urbanik II, T. (1998). "Short-term freeway traffic volume forecasting using radial basis function neural network." *Journal of the Transportation Research Board,* 1651: pp. 39-47.

[198] Park, B. Santra, P. Yun, I. and Lee, D. (2003b). "Development and evaluation of GA-based time-of-day intervals optimization program for traffic signal control." Report, Center for Transportation Studies, University of Virginia, USA.

[199] Park, B. Santra, P. Yun, I. and Lee, D. (2004). "Optimization of time-of-day breakpoints for better traffic signal control." *Transportation Research Record: Journal of the Transportation Research Board,* 1867: 217-223.

[200] Park, D. Rilett, L. R. and Han, G. (1999). "Spectral basis neural networks for real-time travel time forecasting." *Journal of Transportation Engineering.* 125(6), 515-523.

[201] Pauwels, E. J. and Frederix, G. (1999). “Finding salient regions in images: nonparametric clustering for image segmentation and grouping.” *Computer Vision and Image Understanding,* 75(12):73-85.

[202] Payne, H. J. (1971). “Models of freeway traffic and control.” *Simulation council’s proceedings series,* pp. 51-60.

[203] Perrin and Martin. (1998). “On-line comprehensive flow estimation from detectors: model examination.” *1998 annual meeting of TRB, Washington D. C.* Jan 11-15.

[204] Pina, A. C. and Zaverucha, G. (*2006).* “Applying REC analysis to ensembles of sigma-point Kalman filters.” Proc. *Third workshop on ROC analysis in machine learning.*

[205] Polikar, R. (2006). “Ensemble based systems in Decision Making.” *Circuits and Systems Magazine, IEEE,* 6(3):21-45.

[206] Prevedouros, P.D. and Wang, Y. (1999) “Simulation of a Large Freeway/Arterial Network with CORSIM, INTEGRATION, and WATsim”. *Transportation Research Record*, 1678:197-207.

[207] Qi-hua, X. U. Rui, Y. Lianyungang, J. (2005). “Traffic Flow Prediction Using Support Vector Machine Based Method.” *Journal of Highway and Transportation Research and Development.*

[208] Quadstone PARAMICS. (2008). URL: http://www.paramics-online.com/ . Accessed on September 13, 2008.

[209] Quek, C. and Zhou, R. W. (1999). “POPFNN-AARS: A pseudo outer-product based fuzzy neural network.” *IEEE Trans. Syst., Man, Cybern.,* 29(6): 859-870.

[210] Quek, C. Pasquier, M. and Lim, B. B. S. (2006). “POP-TRAFFIC: A novel fuzzy neural approach to road traffic analysis and prediction.” *IEEE transactions on intelligent transportation systems*, 7(2):133-146.

[211] Rahman, S. M. Ratrout, N. T. (2009). Review of the Fuzzy Logic Based Approach in Traffic Signal Control: Prospects in Saudi Arabia. *Article in Press, Journal of Transportation Systems Engineering and Information Technology.*

[212] Rakha, H. A. and Van Aerde, M. W. (1996). “Comparison of Simulation Modules of TRANSYT and Integration Models”. *Transportation Research Record,* 1566:1-7.

[213] Rand, W.M. (1971). “Objective criteria for the evaluation of clustering methods.” *Journal of the American Statistical Association,* 66(336):846-850.

[214] Ratrout, N. T. and Olba, M. A. A. (2009). “Adequacy of TRANSYT-7F and Synchro models along a major arterial in Saudi Arabia.” *Canadian Journal of Civil Engineering,* 36 (1):95–102.

[215] Ratrout, N. T. Rahman, S. M. (2009). “A Comparative Analysis of Currently Used Microscopic and Macroscopic Traffic Simulation Software.” *Article in Press, Arabian Journal for Science and Engineering.*

[216] Roess, R. P. Mc Shane, W. R. and Prassas, E. S. (2004). “Traffic Engineering.” Prentice Hall.

[217] Rousseeuw, P. J. (1987). "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis." *Journal of Computational and Applied Mathematics,* 20(1):53-65.

[218] Sadek, A. W. (2007). "Artificial intelligence applications in transportation." *Transportation Research Circular, Artificial Intelligence in Transportation. Number E-C113, pp. 1-6.*

[219] Sayers, T. M. Bell, M. G. H. Mieden, Th. And Busch, F. (1995). "Improving the traffic responsiveness of signal controllers using fuzzy logic." *The Institution of Engineers,* London, UK.

[220] Sayers, T. M. Bell, M. G. H. Mieden, Th. and Busch, F. (1996). "Traffic responsive signal control using fuzzy logic – a practical modular approach." *The Institution of Engineers,* London, UK.

[221] Shah, A. A. and Dal, L. J. (2007). "Intelligent transportation systems in transitional and developing countries." *IEEE A&E Systems Magazine.*

[222] Shapiro, A. F. (2002). "The merging of neural networks, fuzzy logic, and genetic algorithms." *Insurance: Mathematics and Economics,* 31:115–131.

[223] Shekhar, S. (2004). "Recursive methods for forecasting short-term traffic flow using seasonal ARIMA time series model. "*Ph. D. Dissertation,* Department of Civil Engineering, North Carolina State University.

[224] Shelby, S. G. Bullock, D. M. and Gettman, D. (2005). "Resonant cycles in traffic signal control." *Transportation Research Record: Journal of the Transportation Research Board,* 1925: 215-226.

[225] Smith, B. L, Williams, B. M. and Oswald, K.R. (2002). Comparison of parametric and non-parametric models for traffic flow forecasting. *Transportation Research Part C: Emerging Technologies,* 10(4), 303–21.

[226] Smith, B. L. and Demetsky, M. (1997). “Traffic flow forecasting: Comparison of modeling approaches.” *Journal Transp. Eng.*, 123(4):261-266.

[227] Smith, B. L. and Demetsky, M. J. (1994). “Short-term traffic flow prediction: neural network approach.” *Transportation Research Record,* 1453:98-104.

[228] Smith, B. L. and Demetsky, M. J. (1997). “Traffic flow forecasting: comparison of modeling approaches.” *Journal of Transportation Engineering,* 123(4):261-266.

[229] Smith, B. L. and Oswald, R.K. (2003). “Meeting real-time requirements with imprecise computations: A case study in traffic flow forecasting. *Computer Aided Civil and Infrastructure Engineering,* 18: 201–13.

[230] Smith, B. L. Scherer, W. T. and Hauser, T. A. (2001). “Data-mining tools for the support of signal-timing plan development.” *Transportation Research Record: Journal of the Transportation Research Board,* 1768:141-147.

[231] Smith, B. L. Scherer, W. T. Hauser, T. A. and Park, B. (2002). “Data-driven methodology for signal timing plan development: a computational approach.” *Computer–Aided Civil and Infrastructure Engineering,* 17(6):387-395.

[232] Smith, B. L. William, B. M. and Oswald, R. K. (2002). “Comparison of parametric and nonparametric models for traffic flow forecasting.” *Transportation Research Part C,* 10: 303-321.

[233] Smith, J. O. (2003). "Mathematics of the Discrete Fourier Transform." *W3K publishing,* URL: http:/books.w3k.org.

[234] Smith, S. F. (1980). "A learning system based on genetic adaptive algorithms." *Doctoral dissertation,* Department of computer science. University of Pittsburgh.

[235] Smola, A. and Scholkopf, B. (1997). "On a kernel-based method for pattern recognition, regression, approximation and operator inversion." *Technical Report 1064, GMD.*

[236] Stathopoulos, A. and Karlaftis, M. G. (2003). "A multivariate state space approach for urban traffic flow modeling and prediction." *Transportation Research Part C,* 11:121–135.

[237] Stathopoulos, A. Dimitriou, L. Tsekeris, T. (2008). "Fuzzy Modeling Approach for Combined Forecasting of Urban Traffic Flow." *Computer-Aided Civil and Infrastructure Engineering,* 23(7):521 – 535.

[238] Stephenedes Y. J. Michalopoulos P. G. and Plun R. A. (1981). "Improved estimation of traffic flow for real-time control." *Transportation Research Record,* 795: 28- 39.

[239] Su, H., Zhang, L. and Yu, S. (2007). "Online Support Vector Regression Model for Short-term Traffic Forecasting," *IMSCCS, Second International Multi-Symposiums on Computer and Computational Sciences (IMSCCS 2007),* The University of Iowa, Iowa City, Iowa, USA August 13-August 15, pp.556-559.

[240] Su, M. C. and Chou, C. H. (2001). "A modified version of the K-means algorithm with a distance based on cluster symmetry." *IEEE Trans. Pattern Anal. Machine Intell.,* 23(6):674-680.

[241] Sudkamp, T. Knapp, A. and Knapp, J. (2003). "Model generation by domain refinement and rule reduction." *IEEE transactions on systems, man, and cybernetics-Part B: Cybernetics,* 33(1).

[242] Sun, S. Zhang, C. and Yu, G. (2006). "A Bayesian network approach to traffic flow forecasting. *IEEE Transactions on Intelligent Transportation Systems,* 7(1):124-132.

[243] Sun, S. Zhang, C. and Yu, G. (2006). A Bayesian network approach to traffic flow forecasting. *IEEE Transactions on Intelligent Transportation Systems,* 7(1):124-132.

[244] Takagi, T. and Sugeno, M. (1985). "Fuzzy identification of systems and its application to modeling and control." *IEEE Trans. Syst., Man, Cybern.,* vol. SMC-15, pp. 116-132.

[245] Tan, G. Liu, H. Yuan, W. and Li, C. (2005). "Research of Traffic Flow Forecasting Based on Grids." In *Current Trends in High Performance Computing and Its Applications*, Springer Berlin Heidelberg, pp. 451-456.

[246] Tan, M. -C. Feng, L. -B. and Xu, J. -M. (2007). "Traffic flow prediction based on hybrid ARIMA and ANN model." *China Journal of Highway and Transport., 20*(4):118-121.

[247] Tang, Y. F. Lam, W. H. K. and Ng, P. (2003). “Comparison of four modeling techniques for short-term AADT forecasting in Hong Kong.” *Journal of Transportation Engineering,* ASCE.

[248] Taylor, C. (1994). “Freeway traffic data prediction using artificial neural networks and the development of a fuzzy logic ramp metering algorithm.” *MS thesis,* Department of Electrical Engineering, University of Washington.

[249] Tebaldi, C. West, M. and Karr, A. F. (2002). “Statistical analysis of freeway traffic flows.” *Journal of Forecasting,* 21:39-68.

[250] The MathWorks TM. MATLAB user’s guide. (2010). URL: http://www.mathworks.com/

[251] Thrift, P. (1991). “Fuzzy logic synthesis with genetic algorithms.” *In: Proc. of 4th International Conference on genetic algorithms (ICGA'91),* pp. 509-513.

[252] Topchy, A. Jain, A. and Punch, W. (2003). “Combining multiple weak clusterings.” *Proc. Third IEEE International Conference on Data Mining (ICDM'03),* pp.331-338.

[253] Torgo, L. (2005). “Regression error characteristic surfaces, Proc. *Eleventh ACM SIGKDD int. conf. knowledge discovery in data mining, pp. 697-702, 2005.*

[254] Trafficware Ltd. (2008). URL: http://www.trafficware.com. Accessed on September 13, 2008.

[255] Transportation System Simulation (TSS). (2008). URL: http://www.aimsun.com. Accessed on September 13, 2008.

[256] Tsao, E. C. K. Bezdek, J. C. and Pal, N. R. (1994). “Fuzzy Kohonen clustering networks.” *Pattern recognition*, 27(5):757-764.

[257] Ulbricht, C. (1994). “Multi-recurrent networks for traffic forecasting.” *AAAI-94 Proceedings,* pp. 883-888.

[258] US DoT (2007). ITS unit costs database, Oct. Available at: http://www.itscosts.its.dot.gov.

[259] Van Arem, B. Kirby, H. R. Van Der Vlist, M. J. M. and Whittaker, J. C. (1997). “Recent advances and applications in the field of short-term traffic forecasting.” *International Journal of Forecasting,* 13: 1-12.

[260] Van Arem, B. Whittaker, J. C. and Kirby, H. R. Eds. (1994). *Proceedings of the second DRIVE-II workshop on short-term traffic forecasting*, Report INRO-VVG 1994-19 (TNO-INRO, Delft, Netherlands).

[261] Van der Veen, A. J. Deprettere E. F. and Swindlehurst A. L. (1993). “Subspace-based signal analysis using singular value decomposition.” *Proc. IEEE,* 81:1277-1308.

[262] Vandewalle, J. and De Moor B. (1988). “A variety of applications of singular value decomposition in identification and signal processing, in SVD and Signal Processing: Alogorithms, Applications and Architectures.” E. Deprettere, ed., North-Holland, Amsterdam, pp. 43-91.

[263] Vapnik, V. (1995). “The nature of statistical learning theory.” *Springer Verlag,* New York.

[264] Vapnik, V. and Chervonenkis, A. (1974). “Theory of Pattern Recognition.” Nauka, Moscow, Russia.

[265] Vapnik, V. and Lerner, A. (1963). "Pattern Recognition Using Generalized Portrait Method." *Automat. Remote Contr., 24, 774–780.*

[266] Venturini, G. (1993). "SIA: a supervised inductive algorithm with genetic search for learning attribute based concepts." *In Proc. European Conference on Machine Learning,* pp.280-296, Vienna.

[267] Vlahogianni, E. (2008). "Short-term predictability of traffic flow regimes in signalized arterials." *Road and Transport Research,* 17(2):19-33.

[268] Vlahogianni, E. I. Karlaftis, M. G. and Golias, J. (2005). "Optimized and meta-optimized neural networks for short-term traffic flow prediction: a genetic approach." *Transportation Research Part C: Emerging Technologies,* 13(3):211-234.

[269] Von Altrock, C. (1995). "Fuzzy logic & neuro-fuzzy applications explained." Prentice Hall.

[270] Vonk, E. Jain, L. C. and Johnson, R. P. (1997). "Automatic generation of neural network architecture using evolutionary computation." Advances in fuzzy systems – applications and theory, vol. 14, World Scientific, UK.

[271] Vythoulkas, P. C. (1993). "Alternative approaches to short term forecasting for use in driver information systems." In Transportation and Traffic Theory: *Proceedings of the 12th International Symposium on Transportation and Traffic Theory,* Berkeley, CA, 21-23 July (Elsevier, Amsterdam) pp 485- 506.

[272] Wagenknecht, M. and Hartmann, K. (1988). "Applications of fuzzy sets of type 2 to the solution of fuzzy equation systems." *Fuzzy Sets Syst.,* 25:183-190.

[273] Wang, L. Langari, R. and Yen, J. (1995). “Principal components, B-splines, and fuzzy systems reduction.” In *Proc. CFSA/IFIS/SOFT’95,* Taiwan.

[274] Wang, X. Cottrell, W. and Sichun Mu. (2005). “Using K-means clustering to identify time-of-day breakpoints for traffic signal timing plans**.**” *Proc. 8th International IEEE Conference on Intelligent Transportation Systems.* Vienna, Austria, pp. 586 – 591.

[275] Weijermars, W. and Berkum, E. van. (2005). “Analyzing highway flow patterns using cluster analysis.” *Proceedings of the 8th International IEEE Conference on Intelligent Transportation Systems,* Vienna, Austria, September 13-16.

[276] Whittaker, J. Garside S. and Lindveld, K. (1997). “Tracking and predicting a network traffic process.” *International Journal of Forecasting,* 13(1):51-61, March.

[277] Wiedemann, R. (1974). “Simulation des Straßenverkehrsflusses. Schriftenreihe des.” IfV, 8, Institut für Verkehrswesen, Universität Karlsruhe, (In German language).

[278] Wiedemann, R. (1991). “Modeling of RTI-Elements on Multi-Lane Roads.” In *Advanced Telematics in Road Transport,* The Commission of the European Community, DG XIII, Brussels.

[279] Willaeys, D. and Malvache, N. (1979). “Contributions of the fuzzy sets to man-machine system.” In Advances in Fuzzy set theory and applications, *M. M. Gupta, R. Ragrade and R. Yager, (editors),* North-Holland, Amsterdam, pp. 481-499.

[280] William, B. M. (1999). “Modeling and forecasting vehicular traffic flow as a seasonal stochastic time series process.” *Ph.D. Dissertation,* Department of Civil Engineering, Univ. Virginia, Charlottesville, VA.

[281] Williams, B. M. (2001). “Multivariate vehicular traffic flow prediction: an evaluation of ARIMAX modeling.” *Transportation Research Record,* 1776:194-200.

[282] Wise, B. M. and Ricker, N. L. (1989). “Upset and sensor failure detection in multivariate processes.” Paper 164b, AIChE Meeting, San Francisco.

[283] Wold, S. (1994). “Exponentially Weighted Moving Principal Component Analysis and Projection to Latent Structures.” *Chemo. Intell. Lab. Sys.,* 23:149-161.

[284] Wu, T. Xie, K. Song, G. and Hu, C. (2008). “A multiple SVR approach with time lags for traffic flow prediction.” *Proceedings of the 11th International IEEE Conference on Intelligent Transportation Systems*, Beijing, China, October 12-15.

[285] Xiao, H. Sun, H. Ran, B. and Oh, Y. (2003). Fuzzy-neural network traffic prediction framework with wavelet decomposition. *Transportation Research Record: Journal of the Transportation Research Board,* 1836:16-20.

[286] Xiao, J. -M. and Wang, X. -H. (2004). “Highway traffic flow model using fcm-rbf neural network. *In Advances in Neural Networks,* 317:956-961, Springer Berlin / Heidenlberg.

[287] Xiao-li, Z. and Guo-guang, H. (2007). "Forecasting approach for short-term traffic flow based on Principal Component Analysis and Combined Neural Network." *Systems Engineering – Theory & Practice,* 27(8):167-171.

[288] Xie, Y. and Zhang, Y. (2006). "A Wavelet Network Model for Short-Term Traffic Volume Forecasting." *Journal of Intelligent Transportation Systems,* 10(3):141-150.

[289] Xie, Y. Zhang, Y. and Ye, Z. (2007). "Short-term traffic volume forecasting using kalman filter with discrete wavelet decomposition." *Computer–Aided Civil and Infrastructure Engineering,* 22(5):326-334.

[290] Yager, R. R. (1980). "Fuzzy subsets of type II in decision." *J. Cybern.,* 10:137-159.

[291] Yager, R. R. (1994). "Modeling and formulating fuzzy knowledge bases using neural network." *Neural Netw.,* 7(8):1273-1283.

[292] Yager, R. R. and Filev, D. P. (1994). "Essentials of fuzzy modeling and control." New York: Wiley.

[293] Yam, Y. (1997). "Fuzzy approximation via grid point sampling and singular value decomposition." *IEEE Trans. Syst., Man, Cybern.,* 27:933-951.

[294] Yang, F. M. and Zhu, S. Y. (2006). "Short-term traffic flow forecasting based wavelet." *Journal of Chongqing Jiaotong University,* Chongqing Jiaotong University, Chongqing, pp.99-102.

[295] Yang, Y. (2006). "Spatial-temporal dependency of traffic flow and its implications for short-term traffic forecasting." *Ph. D. Dissertation,* The University of Hong Kong.

[296] Yasdi, R. (1999). “Prediction of road traffic using a neural network approach.” *Neural Computing & Applications,* 8(2):135-142.

[297] Yin, H. Wong, S. C. Xu, J. and Wong, C. K. (2002). “Urban traffic flow prediction using a fuzzy-neural approach.” *Transportation Research Part C: Emerging Technologies,* 10(2): 85-98.

[298] Yin, Y. (2008). Robust Optimal Traffic Signal Timing. *Transportation Research Part B: Methodological,* 42(10): 911-924.

[299] Yu, G. and Zhang, C. (2004). “Switching ARIMA model based forecasting for traffic flow.” *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process,* pp. 429-432.

[300] Yu, G. Q. Hu, J. M. Zhang, C. S. Zhung, L. K., and Song, J. Y. (2003). “Short-term traffic flow forecasting based on Markov chain model.” *In Proc. IEEE Intelligent Vehicles Symp., Columbus, OH,* pp. 208-212.

[301] Yuan, Y. and Zhuang, H. (1996). “A genetic algorithm for generating fuzzy classification rules.” *Fuzzy Sets and Systems.* 84(1):1-19.

[302] Yue, Y. Yeh, A. G.-O. (2008). “Spatiotemporal traffic-flow dependency and short-term traffic forecasting.” *Environment and Planning B: Planning and Design,* 35:762 – 771.

[303] Yun, I. Best, M. and Park, B. (2008). “Evaluation of traffic controller performance during time-of-day transition at coordinated actuated signal system.” *Transportation Research Board 87th Annual Meeting,* No. 08-2176.

[304] Zadeh, L. A. (1973). "Outline of a new approach to the analysis of complex systems and decision processes." *IEEE Trans. Syst., Man, Cybern.,* vol. SMC-3, pp. 28-44.

[305] Zadeh, L. A. (1975). "The concept of a linguistic variable and its application to approximate reasoning", *Information Sciences,* 8:199-249.

[306] Zadeh, L.A. (1965). "Fuzzy Sets, Information and Control."

[307] Zadeh, L.A. (1984). "Making computers think like people." *IEEE Spectrum,* pp. 26-32.

[308] Zhang, D. and Chen, S. (2002). "Fuzzy clustering using Kernel method." *Proceedings of the 2002 International Conference on Control and Automation,* Xiamen, China, June.

[309] Zhang, Q. and Benveniste, A. (1992). "Wavelet networks." *IEEE Trans. Neural Networks,* 3:889-898.

[310] Zhang, Y. and Xie, Y. (2008). "Forecasting of Short-Term Freeway Volume with v-Support Vector." *Transportation Research Record: Journal of the Transportation Research Board,* 2024: 92-99.

[311] Zhang, Y. and Ye, Z. (2008). "Short-term traffic flow forecasting using fuzzy logic system methods." *Journal of Intelligent Transportation Systems,* 12(3):102-112.

[312] Zhao, L. and Wang, F. -Y. (2007). "Short-term fuzzy traffic flow prediction using self-organizing TSK-type fuzzy neural network." *IEEE.*

[313] Zheng, X. and Chu, L. (2008). "Optimal parameter settings for adaptive traffic-actuated signal control." *Proceedings of the 11th International IEEE*

Conference on Intelligent Transportation Systems, Beijing, China, October 12-15, pp. 105-110.

[314] Zhi-Peng, L. Hong, Y. Yun-Cai, L. and Fu-Qiang, L. (2008). "An improved adaptive exponential smoothing model for short-term travel time forecasting of urban arterial street." *Acta Automatica Sinica,* 34(11):1404-1409.

[315] Zhong, M. Sharma, S. and Lingras, P. (2005). "Short-term traffic prediction on different types of roads with genetically designed regression and time delay neural network models." *Journal of Computing in Civil Engineering,* 19(1): 94-103.

[316] Zhong, M. Sharma, S. Lingras, P. (2005). "Short-term traffic prediction on different types of roads with genetically designed regression and time delay neural network models." *Journal of Computing in Civil Engineering,* 19(1):94-103.

APPENDICES

Appendix A- 1: Training data for intersection traffic flow prediction

Ap p1	Ap p2	Ap p3	Ap p4	Ap p5	Ap p6	Ap p7	Ap p8	Ap p9	App 10	App 11	App 12	PAp p5	PAp p6	PAp p7	PAp p8
32	39	50	35	2	3	38	39	20	15	29	23	1	3	31	40
30	35	35	26	1	2	27	33	14	11	22	22	1	2	27	34
25	28	30	30	1	3	24	33	17	12	20	18	1	2	23	31
22	28	27	23	1	2	20	29	12	8	16	19	1	2	21	26
20	24	20	19	1	1	17	22	10	8	13	15	1	1	16	24
13	20	18	16	1	1	14	20	10	5	11	12	0	1	16	20
14	19	14	18	0	2	10	21	7	6	8	14	0	1	12	20
12	19	16	15	0	1	13	17	7	4	10	10	1	1	12	18
12	21	13	17	0	2	10	21	9	6	8	12	0	2	11	22
11	12	11	12	0	1	9	13	6	3	8	8	0	0	11	15
9	10	14	12	0	0	11	13	5	2	10	10	0	0	10	13
6	7	10	7	1	0	8	8	3	2	6	5	1	0	10	11
6	10	8	10	0	1	7	11	4	2	6	7	0	0	8	10
6	9	11	8	0	0	9	10	4	2	8	6	0	0	8	9
6	7	10	8	0	0	8	9	4	2	6	6	0	0	9	7
5	9	11	6	0	0	11	7	4	2	10	4	0	0	11	6
7	7	10	6	0	0	9	7	4	2	7	4	0	0	8	6
6	6	11	6	0	0	9	7	4	2	8	4	0	0	11	7
7	7	14	10	0	0	12	13	4	2	12	13	0	0	11	12
9	14	29	12	1	0	25	14	5	3	24	12	1	1	27	14
10	10	22	10	1	1	18	14	5	3	16	11	0	1	17	16
12	12	29	13	0	2	29	12	6	5	27	8	0	2	25	17
22	20	55	20	1	1	52	20	10	11	46	12	0	1	50	20
33	24	75	30	1	2	79	32	13	14	72	24	1	2	74	29
36	26	82	38	1	2	81	37	15	16	73	27	0	3	75	37
43	31	71	50	1	4	81	49	22	20	71	32	1	2	76	46
65	41	121	64	2	4	125	66	27	25	110	47	2	4	113	62
83	71	171	96	5	7	177	100	35	39	153	80	4	8	167	92
90	63	146	98	3	8	151	93	38	39	128	66	3	7	157	84
111	86	181	112	6	8	195	120	48	48	164	90	5	8	199	113
146	92	224	142	9	9	253	139	57	63	211	107	8	10	272	130
174	134	277	182	16	12	310	182	70	83	254	145	17	13	331	171
124	104	176	134	8	6	196	132	53	57	157	96	6	8	213	133
119	115	183	125	7	10	167	136	57	56	134	95	7	8	168	130
115	113	171	124	7	8	171	131	55	59	133	93	7	7	173	129
123	130	184	136	9	9	184	140	63	69	141	97	8	11	184	138

Ap p1	Ap p2	Ap p3	Ap p4	Ap p5	Ap p6	Ap p7	Ap p8	Ap p9	App 10	App 11	App 12	PAp p5	PAp p6	PAp p7	PAp p8
117	128	165	126	8	9	160	130	64	61	123	84	8	10	160	131
134	138	170	133	10	9	164	139	71	62	128	90	9	10	170	137
150	152	203	146	9	10	176	150	74	65	138	101	10	12	182	138
161	180	217	153	13	12	209	158	81	74	166	105	13	13	202	160
163	174	217	155	12	15	191	156	82	79	149	101	15	13	196	156
174	201	221	154	14	16	204	171	91	82	157	110	13	17	195	166
184	204	247	171	17	15	211	178	96	82	165	113	15	18	218	175
201	227	245	182	18	18	216	198	108	91	166	127	18	21	227	197
209	226	255	188	20	20	224	191	105	98	173	124	20	24	229	195
235	253	262	199	22	23	236	232	127	105	178	148	22	25	237	236
240	253	280	242	29	29	245	254	141	115	186	165	32	31	258	251
258	289	294	244	30	29	269	278	154	123	202	178	31	30	274	279
268	293	314	272	29	33	285	288	156	130	219	187	32	39	301	293
264	314	303	249	26	36	289	291	152	135	221	191	30	38	292	297
276	301	315	250	28	34	282	271	143	129	217	178	30	41	296	270
292	318	311	246	30	35	292	296	159	138	217	190	31	39	310	297
278	308	315	244	27	32	284	268	147	132	214	170	30	35	292	266
270	317	297	225	25	34	266	272	147	128	201	174	25	33	285	275
249	305	302	228	24	32	266	244	132	123	203	158	28	34	274	251
255	321	284	223	25	28	263	271	144	121	202	176	27	32	274	271
248	298	287	225	24	32	243	251	140	119	182	156	24	34	260	258
251	315	274	226	24	31	249	271	147	117	190	170	24	34	262	273
258	287	298	240	24	32	254	263	142	118	196	168	28	35	263	264
253	300	281	239	26	31	253	278	153	119	193	172	28	32	264	287
255	293	300	222	26	33	260	250	142	121	201	157	27	37	261	269
251	306	278	225	23	32	240	274	151	116	184	170	25	35	260	286
265	287	309	249	28	31	253	277	153	115	197	178	30	33	268	278
258	304	303	257	26	33	267	300	164	126	206	188	30	37	272	316
254	306	310	257	31	36	256	295	165	123	198	186	31	39	269	291
253	314	306	253	27	36	266	305	167	125	211	190	30	36	266	308
279	314	355	293	40	42	294	345	197	137	232	219	43	46	288	350
267	329	327	298	35	35	278	356	197	130	217	222	36	42	295	364
281	341	357	318	34	43	306	349	190	147	233	221	37	43	307	350
265	343	325	318	34	39	285	368	196	138	219	231	32	44	286	360
261	307	306	280	27	34	268	307	164	128	209	194	28	37	278	308
263	295	293	244	28	29	260	301	161	120	199	189	28	33	271	302
246	276	298	240	23	30	255	272	147	118	196	168	26	33	261	274
249	286	282	206	24	26	255	266	149	116	195	160	25	32	257	272
240	260	275	202	20	27	238	232	127	114	181	144	25	31	250	249
239	276	265	190	22	27	237	244	133	113	178	152	24	29	254	256

Ap p1	Ap p2	Ap p3	Ap p4	Ap p5	Ap p6	Ap p7	Ap p8	Ap p9	App 10	App 11	App 12	PAp p5	PAp p6	PAp p7	PAp p8
231	259	265	191	21	26	232	226	129	104	181	137	23	31	244	239
225	255	247	169	19	26	221	224	126	104	169	133	21	29	230	238
213	246	245	168	20	29	207	200	119	105	159	118	20	29	219	214
226	248	230	151	22	23	195	219	130	92	152	128	21	28	209	225
209	234	232	167	21	23	196	202	120	91	152	119	20	29	202	213
208	231	216	155	18	25	181	211	120	89	141	126	20	26	182	208
201	208	207	146	18	22	162	191	112	82	127	113	18	27	168	196
193	204	191	129	15	23	155	187	102	81	122	115	15	22	155	201
187	210	203	134	14	24	153	175	95	80	120	109	14	24	152	174
158	184	169	120	12	16	136	171	88	67	104	109	9	15	143	157
140	154	160	115	9	12	121	144	74	53	100	90	9	12	119	136
121	145	144	100	9	10	112	136	68	48	91	86	6	10	105	123
119	136	142	96	6	9	101	119	60	44	80	73	6	8	91	114
105	116	111	84	5	7	92	116	55	35	75	73	4	4	75	97
94	88	99	74	5	5	80	103	51	32	66	63	4	7	71	91
82	88	91	66	3	5	74	86	42	29	60	52	3	5	57	81
71	78	86	55	3	5	68	75	35	26	56	46	3	5	55	74
64	71	75	51	3	5	61	70	33	27	48	42	3	5	51	69
51	64	62	60	2	3	46	69	33	18	39	44	2	3	44	68
46	57	55	38	2	3	42	58	25	18	34	37	2	3	39	54
37	43	43	33	1	3	31	40	19	13	24	24	1	2	28	31
32	34	35	28	1	2	27	34	14	11	21	22	1	2	25	32
26	32	31	28	1	2	23	31	13	9	18	21	1	2	19	31
23	26	28	20	1	2	21	26	12	8	16	16	1	1	17	24
20	23	20	18	1	1	16	24	10	6	13	15	1	1	15	20
17	21	20	16	0	1	16	20	9	5	13	13	0	1	12	18
13	19	16	17	0	1	12	20	9	4	10	12	0	1	11	16
12	18	16	15	1	1	12	18	9	5	10	10	1	1	11	19
13	19	15	16	0	2	11	22	10	5	9	12	1	2	9	18
10	13	14	12	0	0	11	15	6	4	8	10	0	1	10	12
10	12	12	11	0	0	10	13	5	2	9	9	0	0	10	10
7	11	11	9	1	0	10	11	5	3	8	7	0	0	9	9
7	10	10	10	0	0	8	10	4	2	7	6	0	0	7	10
7	9	9	6	0	0	8	9	4	2	7	6	0	0	9	9
5	7	10	7	0	0	9	7	2	2	8	5	0	0	8	7
6	8	12	6	0	0	11	6	2	2	10	5	0	0	8	6
6	7	10	5	0	0	8	6	2	2	7	4	0	0	9	6
6	7	12	6	0	0	11	7	3	2	9	4	0	0	9	7
8	7	12	10	0	0	11	12	3	2	11	12	0	0	10	11
12	10	33	12	1	1	27	14	7	3	26	10	1	1	26	15

Ap p1	Ap p2	Ap p3	Ap p4	Ap p5	Ap p6	Ap p7	Ap p8	Ap p9	App 10	App 11	App 12	PAp p5	PAp p6	PAp p7	PAp p8
10	9	22	11	0	1	17	16	6	3	15	12	1	1	17	14
11	12	26	15	0	2	25	17	7	5	23	11	0	1	23	12
21	19	54	21	0	1	50	20	9	9	45	13	1	1	44	23
33	24	68	29	1	2	74	29	13	14	66	21	1	1	71	31
33	27	79	37	0	3	75	37	17	17	66	23	1	4	70	30
38	29	71	48	1	2	76	46	20	17	66	30	1	2	73	41
60	37	109	64	2	4	113	62	25	23	100	43	1	3	106	59
83	66	150	88	4	8	167	92	33	40	145	72	4	6	161	95
83	61	153	94	3	7	157	84	35	42	131	60	2	9	133	86
107	84	184	112	5	8	199	113	47	48	168	81	5	8	186	103
146	96	234	138	8	10	272	130	58	65	228	93	7	9	260	128
180	129	290	166	17	13	331	171	67	87	271	136	14	12	300	186
117	96	199	133	6	8	213	133	52	67	168	95	8	7	222	139
120	109	174	121	7	8	168	130	56	54	135	90	7	9	179	152
117	114	171	119	7	7	173	129	56	56	137	90	7	9	183	132
124	121	183	132	8	11	184	138	60	71	140	96	8	9	184	141
117	129	168	125	8	10	160	131	65	59	126	85	8	9	156	125
131	140	176	127	9	10	170	137	69	66	131	89	8	9	167	130
137	166	201	140	10	12	182	138	72	67	147	88	8	11	176	140
154	177	211	153	13	13	202	160	84	76	158	105	13	13	194	148
163	181	226	154	15	13	196	156	86	74	155	101	12	11	175	155
173	200	213	160	13	17	195	166	89	83	151	107	12	12	179	154
183	215	247	175	15	18	218	175	93	93	167	113	13	18	203	164
203	237	254	179	18	21	227	197	105	99	174	129	14	17	205	177
206	241	259	192	20	24	229	195	114	103	176	120	16	21	212	193
231	248	267	212	22	25	237	236	124	108	179	154	20	23	222	219
253	260	290	236	32	31	258	251	142	123	193	165	29	26	241	236
261	293	301	243	31	30	274	279	154	126	206	179	26	28	258	263
269	304	334	281	32	39	301	293	160	144	226	190	29	34	278	263
279	324	312	252	30	38	292	297	160	140	221	191	29	35	270	266
285	311	324	255	30	41	296	270	147	145	221	176	27	36	274	244
292	338	332	244	31	39	310	297	157	147	234	197	28	34	279	274
283	317	324	248	30	35	292	266	148	136	222	170	27	32	266	246
276	327	314	230	25	33	285	275	146	134	214	178	25	29	252	246
264	300	312	227	28	34	274	251	144	129	207	157	23	28	251	234
264	322	304	225	27	32	274	271	150	126	211	171	24	28	246	250
260	304	302	226	24	34	260	258	144	124	197	160	20	30	238	244
265	323	301	228	24	34	262	273	147	127	198	172	21	28	226	261
265	299	303	240	28	35	263	264	149	129	200	165	21	29	233	248
276	329	297	246	28	32	264	287	154	123	202	184	23	29	240	269

Ap p1	Ap p2	Ap p3	Ap p4	Ap p5	Ap p6	Ap p7	Ap p8	Ap p9	App 10	App 11	App 12	PAp p5	PAp p6	PAp p7	PAp p8
275	309	309	235	27	37	261	269	154	124	202	166	22	29	248	242
263	324	304	240	25	35	260	286	157	126	199	178	24	25	243	265
280	307	323	256	30	33	268	278	163	125	207	170	25	28	256	274
273	334	316	264	30	37	272	316	172	131	209	200	27	33	262	298
271	310	321	263	31	39	269	291	169	131	207	178	28	37	263	276
273	330	307	253	30	36	266	308	174	130	203	190	27	34	268	301
287	327	360	288	43	46	288	350	201	144	223	222	39	41	275	342
283	353	335	299	36	42	295	364	199	142	229	228	34	40	283	352
287	339	356	318	37	43	307	350	197	148	236	218	33	40	306	332
275	350	333	311	32	44	286	360	191	141	221	227	33	40	284	357
264	329	322	275	28	37	278	308	166	134	214	194	27	36	277	308
276	316	308	240	28	33	271	302	167	130	207	187	27	33	270	301
267	295	298	241	26	33	261	274	154	126	199	165	25	35	261	272
265	306	285	214	25	32	257	272	154	122	196	165	22	30	257	278
254	291	288	214	25	31	250	249	145	122	188	149	20	29	251	251
260	300	284	203	24	29	254	256	144	118	194	156	21	30	252	251
252	282	280	201	23	31	244	239	138	117	187	143	21	28	240	240
240	283	261	180	21	29	230	238	136	112	177	142	22	29	228	239
234	265	255	181	20	29	219	214	124	106	168	127	20	29	209	213
233	268	248	164	21	28	209	225	133	104	160	131	20	29	196	220
235	260	244	169	20	29	202	213	128	102	157	125	21	26	184	208
224	248	227	155	20	26	182	208	121	93	140	124	19	21	166	199
214	222	219	155	18	27	168	196	114	89	130	118	16	21	145	169
198	220	201	135	15	22	155	201	110	79	121	123	15	18	133	174
187	218	213	129	14	24	152	174	96	84	119	108	13	21	134	153
154	185	182	112	9	15	143	157	83	68	111	96	9	14	117	140
139	156	161	107	9	12	119	136	74	53	96	83	8	13	105	122
124	135	137	90	6	10	105	123	61	45	84	78	7	9	94	107
113	122	123	87	6	8	91	114	54	42	70	73	6	8	76	94
103	104	93	68	4	4	75	97	46	29	60	61	5	5	69	81
88	86	90	66	4	7	71	91	44	31	58	57	4	7	60	82
81	79	76	57	3	5	57	81	39	24	46	51	3	5	52	71
80	76	74	55	3	5	55	74	36	24	46	45	3	5	50	67
72	65	66	52	3	5	51	69	33	22	42	42	2	3	42	59
62	60	58	55	2	3	44	68	30	18	35	44	2	3	36	54
49	53	51	41	2	3	39	54	24	15	33	35	2	3	33	44
26	30	38	23	1	2	28	31	15	10	22	19	1	2	27	35
26	27	33	24	1	2	25	32	14	10	19	20	1	2	22	25
20	22	24	29	1	2	19	31	13	7	15	21	1	2	21	30
16	23	21	18	1	1	17	24	11	7	14	16	1	1	16	25

Ap p1	Ap p2	Ap p3	Ap p4	Ap p5	Ap p6	Ap p7	Ap p8	Ap p9	App 10	App 11	App 12	PAp p5	PAp p6	PAp p7	PAp p8
14	19	20	15	1	1	15	20	9	7	12	12	1	1	13	18
13	16	17	14	0	1	12	18	8	5	10	12	1	1	12	15
10	14	15	13	0	1	11	16	7	4	9	10	1	1	12	14
11	14	13	14	1	1	11	19	8	4	9	13	1	1	11	16
11	17	12	13	1	2	9	18	10	4	7	10	1	2	10	21
9	12	13	10	0	1	10	12	5	3	8	8	0	0	9	12
8	9	12	8	0	0	10	10	4	3	8	8	0	1	10	12
7	8	12	7	0	0	9	9	4	3	8	6	0	0	10	8
5	8	9	11	0	0	7	10	5	2	6	5	0	0	7	10
6	8	11	8	0	0	9	9	4	2	8	6	0	1	8	9
5	8	10	7	0	0	8	7	3	2	7	4	0	0	6	8
6	6	9	6	0	0	8	6	3	1	8	4	0	0	10	6
5	7	10	5	0	0	9	6	2	2	7	4	0	0	9	7
5	7	10	6	0	0	9	7	3	2	8	4	0	0	11	7
8	6	10	8	0	0	10	11	3	2	9	10	0	0	9	12
11	8	31	14	1	1	26	15	7	4	23	11	1	1	24	15
10	7	20	9	1	1	17	14	6	4	16	12	1	1	15	12
9	11	24	12	0	1	23	12	5	5	21	8	0	1	22	13
21	16	47	22	1	1	44	23	11	8	40	16	0	1	41	21
33	23	66	26	1	1	71	31	16	14	62	21	1	3	70	24
30	25	73	28	1	4	70	30	13	15	62	20	1	4	74	27
38	27	69	41	1	2	73	41	18	14	65	27	1	2	66	38
62	43	98	59	1	3	106	59	25	22	92	40	2	3	99	48
86	61	145	91	4	6	161	95	32	37	137	75	3	6	156	83
81	59	130	92	2	9	133	86	33	39	111	61	2	10	141	80
106	83	169	104	5	8	186	103	43	47	156	74	6	7	172	100
159	105	228	136	7	9	260	128	56	65	216	91	6	9	237	120
204	136	271	173	14	12	300	186	73	83	244	143	14	13	292	190
147	100	207	136	8	7	222	139	59	65	179	98	8	9	234	142
140	120	184	137	7	9	179	152	60	58	143	111	5	9	179	144
126	120	182	119	7	9	183	132	60	65	141	90	6	10	178	125
126	132	190	130	8	9	184	141	62	68	141	98	6	10	177	146
121	125	168	120	8	9	156	125	63	58	123	80	8	8	154	130
121	139	179	124	8	9	167	130	63	59	133	85	8	7	160	131
142	154	194	131	8	11	176	140	71	66	138	88	8	9	163	129
145	164	200	140	13	13	194	148	78	78	148	97	11	12	180	149
152	170	204	147	12	11	175	155	80	68	135	101	12	12	171	154
166	186	198	145	12	12	179	154	82	69	140	98	12	12	183	150
179	195	233	155	13	18	203	164	88	85	159	105	12	15	195	158
191	212	231	170	14	17	205	177	95	87	158	111	12	18	200	180

Ap p1	Ap p2	Ap p3	Ap p4	Ap p5	Ap p6	Ap p7	Ap p8	Ap p9	App 10	App 11	App 12	PAp p5	PAp p6	PAp p7	PAp p8
202	215	241	187	16	21	212	193	103	93	162	123	16	19	205	195
220	229	250	185	20	23	222	219	117	98	169	141	19	21	215	216
237	245	270	222	29	26	241	236	131	107	184	156	26	26	237	252
244	273	277	221	26	28	258	263	139	118	195	171	26	28	260	263
249	279	304	260	29	34	278	263	144	130	209	171	29	34	281	284
252	303	287	234	29	35	270	266	148	127	208	169	27	31	283	262
258	290	301	231	27	36	274	244	135	129	210	159	25	36	279	257
269	304	307	226	28	34	279	274	144	131	211	180	29	33	282	261
255	292	296	225	27	32	266	246	132	124	202	163	27	35	279	265
255	299	280	208	25	29	252	246	135	119	191	158	26	29	261	242
246	286	288	206	23	28	251	234	130	115	191	147	22	29	252	240
246	294	274	207	24	28	246	250	138	110	190	157	24	29	248	221
244	288	276	219	20	30	238	244	132	111	183	151	22	32	239	241
251	294	259	211	21	28	226	261	141	107	173	161	21	30	234	246
250	270	273	225	21	29	233	248	133	110	177	158	22	30	229	248
250	289	271	218	23	29	240	269	140	114	182	173	19	27	233	260
249	271	294	217	22	29	248	242	132	111	194	153	23	32	248	257
246	289	280	215	24	25	243	265	148	106	189	162	22	27	241	254
269	288	305	251	25	28	256	274	154	112	202	170	24	31	248	274
272	310	296	263	27	33	262	298	163	121	205	186	24	32	274	301
271	296	310	248	28	37	263	276	157	130	200	171	29	36	260	290
257	316	304	246	27	34	268	301	167	125	208	185	28	36	267	301
277	318	342	279	39	41	275	342	191	133	215	219	39	41	286	331
276	347	320	291	34	40	283	352	197	135	220	215	35	41	292	344
283	331	346	302	33	40	306	332	184	143	239	208	36	40	308	348
286	349	322	312	33	40	284	357	193	141	215	223	33	44	298	354
280	312	324	272	27	36	277	308	167	133	212	191	28	37	285	327
279	310	303	238	27	33	270	301	162	127	206	189	27	33	281	294
261	285	298	237	25	35	261	272	153	128	197	164	24	35	266	278
256	302	290	217	22	30	257	278	154	121	194	166	23	32	267	262
262	286	288	212	20	29	251	251	137	121	189	154	21	30	257	249
250	294	273	202	21	30	252	251	139	119	193	152	22	32	250	258
249	277	275	199	21	28	240	240	136	111	184	144	21	31	242	237
235	283	260	175	22	29	228	239	135	113	171	144	18	27	219	224
231	251	253	175	20	29	209	213	124	103	161	128	18	28	199	204
228	253	236	159	20	29	196	220	127	101	150	129	18	25	181	207
222	233	229	160	21	26	184	208	122	95	141	123	16	25	167	187
210	216	212	141	19	21	166	199	116	82	128	118	15	22	154	188
183	200	196	127	16	21	145	169	99	78	110	102	14	22	136	159
171	194	178	115	15	18	133	174	97	70	101	108	12	18	124	164

Ap p1	Ap p2	Ap p3	Ap p4	Ap p5	Ap p6	Ap p7	Ap p8	Ap p9	App 10	App 11	App 12	PAp p5	PAp p6	PAp p7	PAp p8
157	185	181	111	13	21	134	153	85	72	103	95	12	21	127	147
135	163	151	100	9	14	117	140	75	59	89	86	8	15	116	142
121	131	139	87	8	13	105	122	62	51	82	77	8	13	94	120
104	124	120	77	7	9	94	107	54	42	74	68	7	9	87	100
97	97	103	68	6	8	76	94	50	35	60	58	5	8	79	96
80	86	84	58	5	5	69	81	41	28	55	50	4	6	69	88
67	73	77	60	4	7	60	82	43	29	50	50	4	6	62	77
63	69	65	50	3	5	52	71	35	24	42	43	3	4	51	73
57	62	68	47	3	5	50	67	33	22	40	41	2	6	53	71
49	55	54	42	2	3	42	59	28	18	34	36	2	3	44	65
42	47	49	39	2	3	36	54	28	16	28	32	1	3	39	58
35	45	44	31	2	3	33	44	21	13	26	27	2	2	35	47
25	33	35	24	1	2	27	35	13	9	29	37	1	2	27	33
22	26	30	19	1	2	22	25	10	7	26	28	1	2	21	28
19	25	27	28	1	2	21	30	10	7	23	31	1	2	19	28
18	17	21	20	1	1	16	25	8	6	18	25	1	1	17	22
15	17	17	13	1	1	13	18	6	6	15	18	0	1	15	20
12	15	17	13	1	1	12	15	6	5	14	17	1	1	11	17
10	13	16	12	1	1	12	14	5	5	13	15	0	1	11	16
9	15	14	12	1	1	11	16	5	6	11	17	1	1	10	16
11	18	13	15	1	2	10	21	6	8	11	21	1	2	11	20
8	12	12	9	0	0	9	12	5	3	9	12	0	0	10	13
8	10	11	9	0	1	10	12	4	2	11	11	1	1	8	14
6	8	12	7	0	0	10	8	7	2	10	9	1	0	9	9
7	9	9	11	0	0	7	10	3	3	8	10	0	0	8	8
6	7	9	9	0	1	8	9	3	3	8	10	0	0	11	10
5	7	8	8	0	0	6	8	3	2	7	8	0	0	11	8
6	6	11	6	0	0	10	6	3	2	11	7	0	0	20	14
6	6	11	5	0	0	9	7	3	3	9	7	0	1	21	12
5	7	14	6	0	0	11	7	2	2	13	7	0	1	17	11
7	7	11	9	0	0	9	12	3	2	12	8	0	1	18	14
12	10	29	13	1	1	24	15	9	2	29	12	2	2	36	17
9	7	19	11	1	1	15	12	6	2	18	12	2	2	23	16
10	10	22	11	0	1	22	13	9	5	22	14	1	2	28	17
19	15	43	19	0	1	41	21	16	7	39	21	1	2	43	23
29	21	65	26	1	3	70	24	23	9	70	28	2	4	74	29
32	22	74	25	1	4	74	27	22	9	72	30	2	5	76	34
38	28	63	37	1	2	66	38	28	10	72	40	2	4	71	46
59	38	93	44	2	3	99	48	37	11	102	52	2	5	98	57
86	56	141	85	3	6	156	83	55	22	193	81	4	11	157	90

Ap p1	Ap p2	Ap p3	Ap p4	Ap p5	Ap p6	Ap p7	Ap p8	Ap p9	App 10	App 11	App 12	PAp p5	PAp p6	PAp p7	PAp p8
84	57	137	83	2	10	141	80	53	24	165	85	4	11	144	79
100	85	158	99	6	7	172	100	80	33	209	107	7	10	166	97
155	98	205	127	6	9	237	120	117	40	261	131	8	12	249	120
209	130	255	176	14	13	292	190	166	57	341	208	14	16	280	186
147	101	213	138	8	9	234	142	102	48	285	149	11	11	222	139
133	120	182	133	5	9	179	144	106	36	214	147	9	10	172	147
118	113	180	116	6	10	178	125	80	36	210	140	8	11	176	132
125	128	178	140	6	10	177	146	72	42	208	142	10	12	181	141
119	124	166	123	8	8	154	130	58	34	179	146	12	12	156	135
123	130	169	123	8	7	160	131	61	32	181	134	10	12	160	134
133	148	187	125	8	9	163	129	70	38	176	135	10	12	177	143
146	165	193	142	11	12	180	149	77	40	214	158	14	16	190	153
146	164	198	149	12	12	171	154	72	42	192	159	13	15	189	157
164	182	202	139	12	12	183	150	71	42	211	159	14	18	192	157
175	201	228	152	12	15	195	158	75	49	209	173	14	18	201	167
180	219	230	164	12	18	200	180	80	51	227	193	15	19	202	189
204	223	239	178	16	19	205	195	81	58	223	208	19	23	225	194
214	236	246	188	19	21	215	216	87	62	242	244	22	23	223	218
226	247	276	220	26	26	237	252	97	75	271	268	28	28	251	245
239	267	296	233	26	28	260	263	112	85	297	284	27	31	260	259
250	285	308	266	29	34	281	284	112	102	322	295	33	34	269	278
245	299	309	247	27	31	283	262	125	83	320	297	28	33	273	275
262	311	305	227	25	36	279	257	125	87	314	307	28	37	279	256
264	304	316	231	29	33	282	261	135	86	309	299	28	37	290	267
256	290	301	234	27	35	279	265	127	87	308	304	26	37	277	255
244	295	290	215	26	29	261	242	117	79	291	277	25	32	260	257
248	288	289	203	22	29	252	240	105	80	276	277	23	32	245	242
235	300	278	205	24	29	248	221	104	82	263	262	22	33	254	232
237	291	280	206	22	32	239	241	101	79	261	276	25	30	240	236
235	285	281	213	21	30	234	246	96	73	253	280	21	33	234	249
242	282	271	217	22	30	229	248	103	84	248	288	22	31	234	243
257	282	276	220	19	27	233	260	103	80	263	292	23	31	241	269
242	284	286	218	23	32	248	257	102	87	279	297	25	31	239	256
242	298	280	223	22	27	241	254	97	84	271	292	23	29	246	245
262	295	300	248	24	31	248	274	100	98	280	313	26	34	252	279
273	305	316	268	24	32	274	301	123	98	306	337	27	33	267	291
263	304	313	250	29	36	260	290	114	107	299	327	30	37	254	289
260	311	309	254	28	36	267	301	118	108	312	341	26	35	268	299
271	328	352	268	39	41	286	331	120	139	335	392	41	40	271	323
276	349	346	306	35	41	292	344	133	132	353	400	34	41	293	333

Ap p1	Ap p2	Ap p3	Ap p4	Ap p5	Ap p6	Ap p7	Ap p8	Ap p9	App 10	App 11	App 12	PAp p5	PAp p6	PAp p7	PAp p8
282	341	362	303	36	40	308	348	132	158	377	408	34	41	309	328
282	353	345	322	33	44	298	354	137	149	373	422	32	40	297	346
268	322	323	277	28	37	285	327	135	137	352	384	29	34	275	324
270	312	316	253	27	33	281	294	131	117	343	356	26	31	261	281
262	316	305	235	24	35	266	278	128	106	321	343	24	31	263	267
265	298	310	219	23	32	267	262	130	103	311	326	23	29	253	257
250	289	296	210	21	30	257	249	117	98	296	309	20	28	236	242
248	287	291	206	22	32	250	258	116	93	291	307	22	28	227	239
243	282	289	194	21	31	242	237	111	95	279	292	20	30	210	220
229	268	266	175	18	27	219	224	109	90	250	284	19	26	197	206
216	254	249	161	18	28	199	204	96	83	227	264	18	25	187	191
210	243	234	152	18	25	181	207	92	79	212	266	18	23	172	199
201	218	216	146	16	25	167	187	90	79	195	242	17	21	151	185
188	214	198	130	15	22	154	188	80	71	173	241	15	19	144	171
174	185	180	117	14	22	136	159	72	73	153	216	14	19	131	143
151	176	170	110	12	18	124	164	66	71	144	205	13	17	120	144
154	176	178	103	12	21	127	147	70	78	138	206	14	19	119	129
128	158	156	98	8	15	116	142	62	62	132	175	8	15	105	131
116	127	131	87	8	13	94	120	53	52	107	148	7	11	96	110
98	116	116	72	7	9	87	100	40	40	100	129	5	9	85	102
93	100	106	71	5	8	79	96	42	38	89	118	5	9	78	90
84	90	88	62	4	6	69	88	35	35	82	101	5	7	71	86
73	69	78	57	4	6	62	77	32	27	70	90	3	6	59	72
61	66	69	50	3	4	51	73	30	24	58	81	2	5	52	73
62	66	68	49	2	6	53	71	26	25	64	78	3	5	48	63
53	60	61	49	2	3	44	65	22	23	51	68	2	4	45	59
46	59	51	45	1	3	39	58	22	20	47	61	2	3	33	53
39	52	47	36	2	2	35	47	19	20	42	52	1	3	30	46
26	30	36	25	1	2	27	33	15	12	30	36	2	3	27	32
24	29	28	22	1	2	21	28	12	12	24	31	2	3	21	30
20	24	26	25	1	2	19	28	10	8	23	28	1	3	17	27
17	21	22	17	1	1	17	22	9	7	20	24	1	2	15	24
13	19	19	17	0	1	15	20	9	7	17	20	1	3	14	20
12	17	15	15	1	1	11	17	6	5	12	17	1	2	15	19
10	16	14	14	0	1	11	16	6	6	13	17	1	1	11	19
10	16	13	12	1	1	10	16	5	6	14	16	1	1	12	20
11	20	15	15	1	2	11	20	6	9	12	21	1	3	9	22
8	12	13	10	0	0	10	13	4	4	11	13	1	2	9	15
7	9	11	10	1	1	8	14	5	2	9	11	1	1	7	17
6	9	11	9	1	0	9	9	6	3	9	9	0	0	11	7

Ap p1	Ap p2	Ap p3	Ap p4	Ap p5	Ap p6	Ap p7	Ap p8	Ap p9	App 10	App 11	App 12	PAp p5	PAp p6	PAp p7	PAp p8
6	8	10	9	0	0	8	8	3	3	9	9	0	1	6	11
8	10	12	8	0	0	11	10	4	3	12	9	0	1	8	10
7	8	11	12	0	0	11	8	5	2	11	8	0	0	7	7
10	9	15	12	0	0	20	14	5	3	21	9	0	0	10	12
9	13	21	10	0	1	21	12	5	3	21	9	0	0	13	9
6	14	23	11	0	1	17	11	3	2	19	10	0	0	12	8
10	16	26	11	0	1	18	14	7	3	21	11	0	1	11	11
14	23	45	17	2	2	36	17	11	3	41	13	2	2	27	18
15	17	34	18	2	2	23	16	7	3	26	14	1	2	19	14
21	21	33	20	1	2	28	17	10	6	28	17	0	1	26	16
32	26	47	27	1	2	43	23	18	6	44	25	1	1	42	20
39	30	74	33	2	4	74	29	22	10	75	31	2	3	72	24
41	31	76	36	2	5	76	34	20	11	73	33	2	6	76	24
49	39	71	49	2	4	71	46	28	11	76	45	1	6	65	39
67	49	96	57	2	5	98	57	36	15	101	60	2	7	100	48
89	65	147	90	4	11	157	90	57	25	192	87	5	11	144	84
84	68	141	84	4	11	144	79	53	26	158	88	5	15	135	81
102	93	154	95	7	10	166	97	72	30	210	106	6	13	162	93
156	108	220	128	8	12	249	120	120	42	271	137	9	14	241	113
195	139	250	178	14	16	280	186	157	53	330	199	14	19	283	181
148	108	208	137	11	11	222	139	99	45	274	153	12	15	220	142
142	124	181	137	9	10	172	147	106	40	202	150	9	13	187	136
124	122	177	123	8	11	176	132	83	38	208	140	8	12	169	130
131	140	186	138	10	12	181	141	75	41	214	139	10	13	181	142
128	134	170	126	12	12	156	135	63	36	178	147	9	13	155	129
133	142	172	129	10	12	160	134	66	36	181	141	12	13	157	143
146	159	194	133	10	12	177	143	74	41	187	152	13	16	167	134
159	177	202	148	14	16	190	153	78	43	220	162	16	21	188	142
156	179	218	157	13	15	189	157	78	45	208	162	21	19	185	147
174	203	210	148	14	18	192	157	80	47	217	166	20	22	194	149
188	204	236	164	14	18	201	167	84	49	213	178	20	24	195	169
196	235	236	173	15	19	202	189	76	53	234	202	23	28	216	183
207	228	257	184	19	23	225	194	86	61	243	207	25	31	210	207
220	249	262	199	22	23	223	218	88	66	256	243	30	32	232	207
234	253	280	220	28	28	251	245	99	77	275	259	37	42	244	245
249	277	306	233	27	31	260	259	110	86	298	280	36	36	265	276
251	292	307	259	33	34	269	278	119	98	317	293	40	44	267	290
252	305	302	248	28	33	273	275	127	89	311	306	38	42	274	281
260	317	308	234	28	37	279	256	134	87	312	307	42	49	281	274
275	306	312	243	28	37	290	267	142	89	314	308	38	48	305	288

Ap p1	Ap p2	Ap p3	Ap p4	Ap p5	Ap p6	Ap p7	Ap p8	Ap p9	App 10	App 11	App 12	PAp p5	PAp p6	PAp p7	PAp p8
259	294	318	231	26	37	277	255	132	88	296	294	33	47	281	265
247	299	302	219	25	32	260	257	124	77	279	291	36	45	275	265
252	290	289	213	23	32	245	242	114	81	268	277	33	48	251	260
244	294	292	212	22	33	254	232	109	81	272	273	35	46	262	252
250	294	279	216	25	30	240	236	102	77	260	272	34	47	247	251
248	285	284	220	21	33	234	249	103	83	246	284	33	45	247	265
238	274	283	219	22	31	234	243	105	85	249	287	33	44	232	260
250	294	284	226	23	31	241	269	100	84	265	304	34	45	241	293
255	284	283	219	25	31	239	256	106	85	268	296	32	44	242	263
247	296	285	222	23	29	246	245	100	86	274	285	31	48	251	269
262	293	303	251	26	34	252	279	110	99	282	312	39	50	257	273
268	311	319	263	27	33	267	291	124	101	294	326	36	48	274	296
259	303	315	246	30	37	254	289	119	104	290	334	40	47	265	291
255	306	310	251	26	35	268	299	119	107	308	342	40	54	265	315
265	323	339	265	41	40	271	323	126	135	321	383	47	59	273	344
279	342	333	287	34	41	293	333	131	130	354	387	50	56	308	361
264	339	366	290	34	41	309	328	135	150	365	401	46	54	300	344
263	343	355	323	32	40	297	346	135	147	365	408	42	51	308	357
268	314	332	270	29	34	275	324	133	132	327	369	44	51	284	337
255	312	307	241	26	31	261	281	129	114	326	345	38	47	265	291
250	297	299	228	24	31	263	267	133	112	305	328	39	46	260	272
250	285	300	218	23	29	253	257	126	98	292	316	36	41	256	270
243	276	279	194	20	28	236	242	120	93	267	292	32	42	245	242
232	270	270	191	22	28	227	239	118	96	263	290	32	42	233	241
230	270	258	179	20	30	210	220	111	98	247	271	30	43	216	235
213	248	232	163	19	26	197	206	101	85	231	259	29	46	201	221
201	229	228	156	18	25	187	191	96	77	206	244	30	41	191	202
194	225	214	147	18	23	172	199	91	77	191	239	30	37	181	199
189	217	207	141	17	21	151	185	90	70	170	234	29	36	162	188
176	192	188	128	15	19	144	171	80	68	164	225	23	38	153	187
156	174	177	110	14	19	131	143	74	66	149	199	22	30	129	161
156	159	157	102	13	17	120	144	69	64	135	190	21	31	127	156
140	160	166	98	14	19	119	129	73	78	127	186	19	28	118	142
132	148	145	89	8	15	105	131	63	65	123	171	15	21	109	139
110	124	129	80	7	11	96	110	55	54	104	145	12	19	92	115
96	114	115	72	5	9	85	102	42	43	102	127	9	14	82	98
87	91	107	67	5	9	78	90	44	40	87	109	7	18	72	90
80	86	92	60	5	7	71	86	37	34	83	101	7	12	64	85
73	73	77	54	3	6	59	72	39	39	64	87	5	9	55	77
60	69	70	52	2	5	52	73	32	28	61	82	5	9	49	73

Ap p1	Ap p2	Ap p3	Ap p4	Ap p5	Ap p6	Ap p7	Ap p8	Ap p9	App 10	App 11	App 12	PAp p5	PAp p6	PAp p7	PAp p8
60	61	65	45	3	5	48	63	26	26	53	73	3	11	43	66
53	54	58	42	2	4	45	59	26	21	53	61	3	6	39	62
47	48	45	38	2	3	33	53	20	23	40	61	3	5	35	53
39	44	43	32	1	3	30	46	17	19	37	49	3	5	30	47

Appendix A- 2: Testing data for intersection traffic flow prediction

Ap p1	Ap p2	Ap p3	Ap p4	Ap p5	Ap p6	Ap p7	Ap p8	Ap p9	App 10	App 11	App 12	PAp p5	PAp p6	PAp p7	PAp p8
23	26	35	23	1	2	23	30	12	11	27	32	1	2	26	33
21	24	27	22	1	2	20	27	12	7	23	29	1	3	19	27
17	20	23	20	1	2	15	21	7	7	19	22	1	2	16	23
15	19	21	14	1	1	14	19	6	7	16	20	1	1	14	19
12	16	17	15	1	1	12	17	6	6	14	18	1	1	13	18
11	16	13	12	0	1	9	16	5	5	9	15	0	1	12	17
10	15	14	9	1	1	9	15	5	5	11	15	0	1	9	15
9	17	12	12	0	1	8	16	4	5	10	14	0	1	9	17
10	16	15	15	0	1	10	20	5	8	11	20	0	2	9	22
8	11	10	9	0	0	7	12	6	3	8	13	0	1	9	13
7	7	11	10	0	0	8	13	5	3	9	9	0	0	10	15
5	8	11	7	1	0	8	8	5	2	9	8	1	0	8	9
5	7	9	6	0	0	6	8	2	2	6	8	0	1	7	8
6	8	11	9	0	0	8	10	3	3	8	11	0	0	8	9
5	6	8	8	0	1	6	7	3	2	6	8	0	0	8	10
6	6	14	9	0	0	10	11	4	2	11	10	0	0	11	11
5	7	15	7	0	1	12	8	4	3	12	7	0	0	12	10
6	6	12	8	0	1	9	9	2	2	11	8	0	0	11	9
7	7	12	11	1	1	11	11	6	2	11	9	1	1	13	13
10	11	31	14	1	1	24	15	10	3	28	11	2	1	22	15
9	7	20	10	1	1	14	12	9	3	16	12	1	2	15	13
11	12	26	10	1	1	22	14	10	3	23	13	0	1	19	14
21	19	43	16	0	0	40	17	18	6	41	17	1	1	38	16
28	23	75	20	1	3	72	23	24	8	70	23	1	3	78	25
28	20	78	26	1	3	74	26	19	8	71	27	2	3	69	28
39	28	54	35	2	3	54	37	24	9	63	34	1	4	61	31
55	36	89	43	2	4	86	42	39	13	94	44	2	4	93	43
90	59	140	67	4	8	142	70	62	23	174	68	5	7	138	72
72	57	132	72	4	8	131	70	55	25	154	74	4	9	130	69
101	84	158	86	5	8	160	88	86	31	204	89	5	7	162	82
161	99	226	113	8	10	243	103	130	39	273	115	7	12	241	111
193	132	258	168	15	13	282	176	179	57	343	185	15	16	275	170
137	106	200	117	10	12	215	127	111	47	269	133	10	13	219	134
132	100	183	116	7	10	170	127	116	33	201	128	8	9	175	128
114	106	169	103	6	9	159	110	94	36	184	121	7	10	157	116
122	118	176	114	10	10	164	124	73	37	199	126	9	12	161	124
117	108	146	99	7	8	132	108	58	29	151	118	9	9	139	108

App1	App2	App3	App4	App5	App6	App7	App8	App9	App10	App11	App12	PApp5	PApp6	PApp7	PApp8
114	120	153	103	7	8	128	107	59	31	142	107	7	11	134	116
123	134	164	104	8	9	143	110	65	36	158	119	8	11	143	115
130	149	181	118	10	11	154	121	75	38	173	129	12	13	160	130
138	154	168	117	11	11	145	128	62	35	165	132	12	13	155	135
152	160	189	126	11	13	153	129	70	37	169	137	11	15	163	141
159	174	197	126	9	14	167	135	67	44	183	144	12	17	174	150
162	195	224	152	13	18	177	150	73	46	196	161	13	20	193	156
174	202	209	144	13	19	172	165	76	51	198	175	14	19	177	179
195	209	239	168	17	19	197	181	82	58	209	211	17	23	203	192
201	217	266	193	23	24	211	214	91	67	239	229	24	28	221	230
211	243	282	208	23	27	227	234	107	79	259	247	24	28	237	247
220	264	290	242	27	31	242	268	114	90	284	275	27	32	251	276
228	274	291	222	27	33	242	240	115	78	287	271	26	34	251	247
233	287	298	204	25	31	259	231	129	72	283	268	25	35	266	231
246	282	320	221	24	35	282	242	137	81	303	273	25	36	273	251
236	277	309	215	23	30	248	234	129	83	272	267	23	35	264	243
237	269	303	198	20	26	243	231	129	75	253	262	23	31	247	235
230	261	283	189	20	27	223	220	110	74	244	256	21	33	229	226
222	268	278	193	21	29	229	214	106	71	252	252	22	31	230	221
228	269	280	189	21	30	222	215	99	67	234	237	20	30	226	223
221	264	274	200	21	27	218	227	102	75	230	254	20	30	217	223
218	254	277	201	21	26	210	219	103	77	230	270	22	32	217	233
234	260	276	207	19	26	213	250	104	80	236	277	21	30	221	250
229	268	277	190	22	29	221	226	102	80	258	269	20	33	232	236
225	280	272	198	21	27	219	227	95	77	245	258	20	29	224	229
239	272	293	227	24	30	229	257	104	95	258	288	24	32	232	260
248	278	307	241	23	29	244	261	131	100	267	298	25	33	242	271
245	277	304	218	29	34	235	263	123	104	270	305	27	40	237	270
239	292	297	233	27	34	243	277	126	104	284	324	27	36	254	286
250	299	346	239	36	43	260	315	129	137	308	372	38	41	259	324
249	320	345	271	36	41	280	319	133	133	337	373	34	42	283	334
253	331	350	283	33	41	285	330	143	151	342	393	32	44	296	344
246	336	365	311	33	41	280	343	144	153	346	401	31	42	287	348
239	306	323	251	26	36	252	300	138	129	314	372	25	40	265	318
236	280	296	220	25	31	232	265	125	108	295	330	24	35	251	276
235	277	289	213	22	32	233	256	120	106	282	311	23	35	244	267
232	263	281	199	21	29	230	246	130	98	262	292	21	34	242	251
215	243	279	175	18	27	212	211	116	88	242	262	20	31	217	230
213	239	272	166	19	26	211	214	112	89	243	260	19	30	219	228
208	231	242	160	18	24	185	198	108	88	216	249	19	29	208	208

Ap p1	Ap p2	Ap p3	Ap p4	Ap p5	Ap p6	Ap p7	Ap p8	Ap p9	App 10	App 11	App 12	PAp p5	PAp p6	PAp p7	PAp p8
198	212	222	134	17	23	178	184	98	78	207	229	17	30	189	203
179	201	214	135	18	22	164	169	92	71	188	216	16	26	172	183
174	196	208	120	18	23	155	166	85	75	172	206	17	27	163	176
160	182	193	123	13	22	142	155	82	67	159	198	14	25	147	164
158	175	172	102	13	19	128	146	74	66	146	187	14	23	132	154
144	154	161	97	13	19	115	131	69	59	131	179	12	23	123	134
135	141	148	88	11	18	109	131	61	59	126	163	11	22	119	129
130	146	153	87	11	20	106	121	65	73	121	172	12	23	115	127
122	128	150	80	8	14	101	116	57	69	115	153	9	18	107	121
103	103	119	77	6	13	84	103	50	54	98	130	7	13	90	110
87	95	111	58	6	9	78	87	39	41	90	109	6	11	82	96
73	77	95	55	6	7	64	76	37	36	73	99	6	11	72	88
68	71	78	54	3	6	60	73	34	34	69	82	5	8	62	84
57	61	77	51	3	6	53	68	31	26	60	75	3	8	55	72
53	56	63	43	3	5	44	61	26	21	52	66	3	6	50	66
49	47	62	39	2	5	43	52	25	20	51	57	3	6	46	58
42	44	52	35	1	3	37	49	22	20	44	52	3	4	42	53
33	43	44	37	2	3	33	47	18	22	38	50	2	3	33	50
32	34	41	31	2	2	28	39	17	15	35	45	1	3	33	46

Appendix A- 3: Steps to build abductive network model

The Abductive Model building is achieved by using Modelquest Prospector developed by AbTech Corporation. The following basic steps are to be followed to build the model.

1. The data including both input and output should be saved as .txt file.
2. The data in txt format should be imported in the Modelquest environment.
3. A new model is to be built by selecting input and output variables. At the same time, an index is to be provided which will be maintained internally by the software.
4. The value of *CPM* can also be provided if the model building window is active. The user can also provide some optional parameters such as number of layers and number of input at the first layer.
5. The *data generation* menu is to be selected to mention the % of training data. In this stage, it is also possible to provide the nature of the data whether sequential or random. The training and evaluation dataset will be generated.
6. The model *synthesis* menu is to be selected to build the network.
7. The *evaluation* menu will help to evaluate the testing datasets. It will provide some error measures such as Mean Absolute Error (MAE) and R-square value for training and testing data.
8. The user can export the estimated data in txt format for further analysis.
9. The user can get the co-efficient of the model by selecting the *encode* menu. The encoded model can be used with other model for example in Microsoft Excel^{TM}.
10. A typical screenshot of the software is shown below.

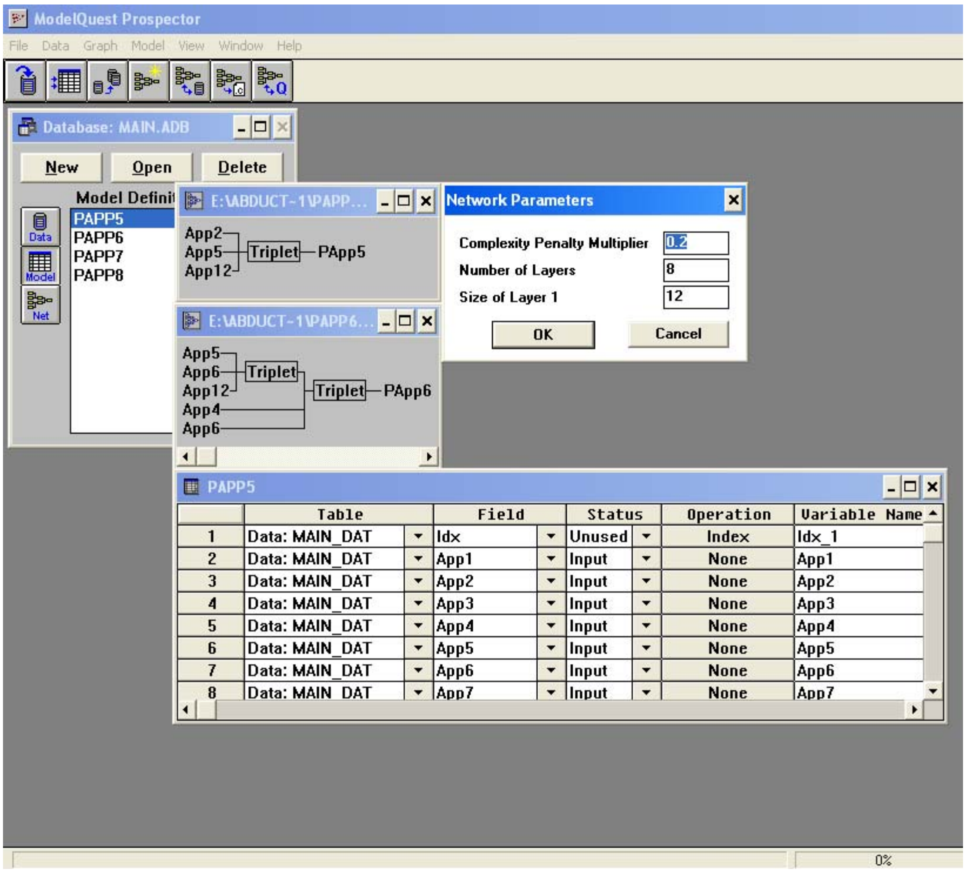


Figure 1: A screenshot of the Modelquest showing few windows of model, network and network parameters.

Appendix A- 4: Steps to build ANFIS model

The ANFIS model is built in MATLAB environment by using available functions. This software provides the required functions to perform normalization, clustering, fuzzy inference system (FIS) building, error analysis along with other functions to perform mathematical and computational analysis. The required steps to build ANFIS model are as follows:

1. The first step is to bring the input and output data in MATLAB^{TM} environment by using *load* function.
2. The data are normalized between -1 to +1 by using the function named *mapminmax*.
3. The initial FIS is obtained with the help of function named *genfis3*. This function uses Fuzzy C-Means clustering algorithm to build the FIS. The training data is used to build the FIS.
4. The parameters of the FIS are learnt by using the function known as *anfis*. The initial FIS is provided in this function as input.
5. The testing inputs are used to find out the predicted output.
6. Different error measures such as Root Mean Squared Error (RMSE), Mean Absolute Error (MAE) are to be calculated to investigate the performance of the developed models.
7. A sample code is shown in the following table.

Table 1: A sample code to implement ANFIS model.

```
clear all;
j=5;
 %% Input & Output Selection
t_in_out=load('main.txt');
 [t_row t_col]=size(t_in_out);
t_months=t_row/96;
n_tr_months=6; % Mention No. of months data for training.
n_tst_months=2;
n_tr_data=96*n_tr_months; % Number of 15-min training intervals.
 tr_data=t_in_out(1:n_tr_data,:); %Training Data.
%%Input & Output Data for Int-2
tr_in=tr_data(1:n_tr_data-96,3:7); %Input Training Data for
% forecasting 15-min traffic vol of Int-2.
tr_out=tr_data(97:n_tr_months*96,:);
 n_tst_data=96*n_tst_months; % Number of 15-min testing intervals.
tst_data=t_in_out(n_tr_data+1:n_tr_months*96+n_tst_data,:);
tst_in=tst_data(1:n_tst_data-96,3:7);
tst_out=tst_data(97:end,:);

    tst_in=tst_in';

        [tr_in,ps_tr_in] = mapminmax(tr_in);tr_in=tr_in'; tr_in=(tr_in+1)/2;
        tst_in=mapminmax('apply',tst_in,ps_tr_in);tst_in=tst_in';tst_in=(tst_in+1)/2;
 %%  ANFIS Model Development-FCM
rand('state',123);randn('state',234);
%for Southbound of Int-2
%Initial FIS by GENFIS3
 tr_initial_fis=genfis3(tr_in, tr_out(:,j),'sugeno',3,[2,400,1e-5,1]);
%Initial FIS  by GENFIS2
% tr_initial_fis=genfis2(tr_in, tr_out(:,j),0.2);
% Initial FIS by GENFIS1
%  tr_initial_fis=genfis1([tr_in tr_out(:,j)],3);
epoch_n=8;

[tr_anfis, error1,ss,tst_anfis,error2]=...
anfis([tr_in tr_out(:,j)],tr_initial_fis,[epoch_n 0 0.01 0.9 1.1],[],[tst_in tst_out(:,j)],1);
tst_out_anfis1=evalfis(tst_in,tr_anfis);
tr_out_anfis1=evalfis(tr_in,tr_anfis);
% end
% figure;
% plot(tst_out(:,j),'-r');hold on;plot(tst_out_anfis1(:,j),'-b');
figure;plot([error1 error2]);

%% %%  ANFIS Model Development -model SUBTRACT
rand('state',223);randn('state',334);
%for Southbound of Int-2
%Initial FIS by GENFIS3
%  tr_initial_fis=genfis3(tr_in, tr_out(:,j),'sugeno',3,[2,400,1e-5,1]);
%Initial FIS  by GENFIS2
tr_initial_fis2=genfis2(tr_in, tr_out(:,j),0.5);
% Initial FIS by GENFIS1
%  tr_initial_fis=genfis1([tr_in tr_out(:,j)],3);
epoch_n=100;
```

```
[tr_anfis2, error1,ss,tst_anfis2,error2]=...
anfis([tr_in tr_out(:,j)],tr_initial_fis2,[epoch_n 0 0.01 0.9 1.1],[],[tst_in tst_out(:,j)],1);
tst_out_anfis2=evalfis(tst_in,tr_anfis2);
tr_out_anfis2=evalfis(tr_in,tr_anfis2);
% end
% figure;
% plot(tst_out(:,j),'-r');hold on;plot(tst_out_anfis1(:,j),'-b');
figure;plot([error1 error2]);

%% NN MODEL DEVELOPMENT

% preparing training data
tr_input=tr_data(1:n_tr_data-96,3:7); % Input Data Selection
tr_input=tr_input';
[n_tr_input, n_s_input]=mapstd(tr_input);
tr_output=tr_data(97:end,:);% output data
tr_output=tr_output(:,j)';
[n_tr_output, n_s_output]=mapstd(tr_output);

%preparing testing data
test_input=tst_data(1:n_tst_data-96,3:7);
test_input=test_input';
test_output=tst_data(97:end,:);
test_output=test_output(:,j)';
n_test_input=mapstd('apply',test_input,n_s_input);
n_test_output=mapstd('apply',test_output,n_s_output);
%% Netowrk Building
rand('state',54);randn('state',23);
net = newff(n_tr_input,n_tr_output(1,:),[4]); % 1 is used because we are getting one
% output at a time
     net.divideFcn = '';
     net.trainParam.show = 50;
     net.trainParam.lr = 0.005;
     net.trainParam.epochs =100;
     net.trainParam.goal = 1e-4;
 [net,tr]=train(net,n_tr_input(:,:),n_tr_output(1,:));
   n_tr_output_model = sim(net,n_tr_input);

   tr_output_model=mapstd('reverse',n_tr_output_model,n_s_output);
%    plot(tr_output_model,'-r');hold on; plot(tr_output(1,:),'-b');
  tr_mae= mae(tr_output(1,:)-(tr_output_model));



  %testing
  n_test_output_model=sim(net,n_test_input);
  test_output_model=mapstd('reverse',n_test_output_model,n_s_output);
%    figure
%    plot(test_output_model,'-r');hold on; plot(test_output(1,:),'-b');
  v_mae_nn_test=mae(test_output(1,:)-test_output_model);

%   plotperform(tr)

 %% error plotting
```

```
X1=tst_out_anfis1;X2=test_output_model';X3=tst_out_anfis2;
 y=tst_out(:,j);[endy c]=size(y);
%   for i=1:endy; y==0;y(find(y==0))=1;end;

% MaxErr=abs(Y-y);
figure; plot(0.25:0.25:24,y,'-'); legend('Actual');
hold on; plot(0.25:0.25:24,test_output_model,'o');legend('Actual','NN Output');
hold on;plot(0.25:0.25:24,tst_out_anfis1,'s');legend('Actual','NN','ANFIS-FCM');
hold on;plot(0.25:0.25:24,tst_out_anfis2,'^');legend('Actual','NN','ANFIS-FCM','ANFIS-SC');
xlabel('Time (Hour)');ylabel('Traffic Volume (veh/15min)');axis tight;

v_rsquare_nn=rsquare(y,X2);
v_rmse_nn=errperf(y,X2,'rmse');%v_mape_nn=errperf(y,X2,'mape');

% % For ANFIS Model
%v_mare_anfis=errperf(y,tst_out_anfis1,'mare');
v_rsquare_anfis=rsquare(y,tst_out_anfis1);
v_rmse_anfis1=errperf(y,tst_out_anfis1,'rmse');%v_mape_anfis=errperf(y,tst_out_anfis1,'mape');
v_mae_anfis1=errperf(y,tst_out_anfis1,'mae');
v_rsquare_anfis2=rsquare(y,tst_out_anfis2);
v_rmse_anfis2=errperf(y,tst_out_anfis2,'rmse');
v_mae_anfis2=errperf(y,tst_out_anfis2,'mae');

mean_err_anfis1=mean(y-tst_out_anfis1);std_err_anfis1=std(y-tst_out_anfis1);
mean_err_anfis2=mean(y-tst_out_anfis2);std_err_anfis2=std(y-tst_out_anfis2);
mean_err_nn=mean(y-test_output_model');std_err_nn=std(y-test_output_model');

fis_fcm=getfis(tr_anfis);
ruleview(tr_anfis)
figure;plotfis(tr_anfis)
figure;gensurf(tr_anfis)
ruleedit(tr_anfis)
```

Appendix A- 5: Steps to build type-2 fuzzy logic model

The type-2 FL model is built in MATLAB™ environment. But the required computer codes to perform type-2 fuzzy operations are obtained from the website of Professor Mendel who is the leading scholar in the area of type-2 FL model building. The interested readers can download the codes from the URL: http://sipi.usc.edu/~mendel/software/. The required steps to build proposed type-2 FL model are as follows.

1. The first step is to bring the input and output data in MATLAB™ environment by using *load* function.
2. The second step is to prepare the training and testing data.
3. The user has to provide "mean" and "standard deviation" of Gaussian functions of the membership functions to build type-1 non-singleton fuzzy inference system.
4. The singular value decomposition can be used to reduce number of rules. The function named *svd_qr_sfls_type1* should be used to perform this action.
5. The type-1 model is to be trained with the reduced rules and training data. The user can use the function called *train_sfls_type1*.
6. The type-2 FL model is to be trained using the initial values of "mean" and "sigma" obtained by using type-1 FL model. The user can use the function called *train_sfls_type2*.
7. The user can obtain the predicted values by using the testing input and the type-2 FL model. The function named *nsfls2* can be used to perform the mentioned action.

8. Different error measures such as RMSE, MAE are to be calculated to investigate the performance of the developed models.

Appendix A- 6: Description of a few error measurement functions

In this dissertation, a few error measurement functions of MATLAB™ are used which are described below. The interested readers can consult the help of MATLAB™.

1. Mean squared error performance function

Syntax

perf = mse(E,Y,X,FP)

dPerf_dy = mse('dy',E,Y,X,perf,FP)

dPerf_dx = mse('dx',E,Y,X,perf,FP)

info = mse(code)

Description

mse is a network performance function. It measures the network's performance according to the mean of squared errors.

mse(E,Y,X,FP) takes E and optional function parameters,

E	Matrix or cell array of error vectors
Y	Matrix or cell array of output vectors (ignored)
X	Vector of all weight and bias values (ignored)
FP	Function parameters (ignored)

and returns the mean squared error.

mse('dy',E,Y,X,perf,FP) returns the derivative of perf with respect to Y.

mse('dx',E,Y,X,perf,FP) returns the derivative of perf with respect to X.

mse('name') returns the name of this function.

mse('pnames') returns the names of the training parameters.

mse('pdefaults') returns the default function parameters.

2. Mean absolute error performance function

Syntax

perf = mae(E,Y,X,FP)

dPerf_dy = mae('dy',E,Y,X,perf,FP)

dPerf_dx = mae('dx',E,Y,X,perf,FP)

info = mae(code)

Description

mae is a network performance function. It measures network performance as the mean of absolute errors.

mae(E,Y,X,FP) takes E and optional function parameters,

E Matrix or cell array of error vectors

Y Matrix or cell array of output vectors (ignored)

X Vector of all weight and bias values (ignored)

FP Function parameters (ignored)

and returns the mean absolute error.

mae('dy',E,Y,X,[perf,FP) returns the derivative of perf with respect to Y.

mae('dx',E,Y,X,perf,FP) returns the derivative of perf with respect to X.

mae('name') returns the name of this function.

mae('pnames') returns the names of the training parameters.

mae('pdefaults') returns the default function parameters.

Appendix A- 7: The training and testing data for the freeway traffic flow model.

St125	St109	St303	St311	St313	St315	St401	St801	St803	St805	St1201
749	447	613	44	461	3172	88	103	112	195	46
645	322	507	33	394	2490	65	91	87	139	34
492	257	414	25	320	1933	47	79	68	99	25
375	219	349	25	267	1590	40	66	60	96	19
307	211	340	33	248	1425	44	69	63	107	16
332	221	409	42	296	1400	84	107	73	241	18
466	341	635	97	409	2192	153	147	142	343	36
482	475	828	102	482	2889	148	172	164	227	62
464	578	723	115	498	3006	153	164	158	208	76
475	600	716	125	515	3331	159	163	174	287	82
551	592	694	134	541	3534	165	166	189	286	82
673	619	747	129	557	3736	174	179	198	292	84
685	583	804	124	523	3708	163	169	187	243	87
711	574	727	150	473	3554	144	159	201	221	90
908	663	800	129	531	4352	168	176	215	266	106
988	710	963	125	618	4381	178	185	222	258	119
1051	785	964	171	646	4148	175	410	233	262	116
1004	874	1098	180	699	4972	186	308	236	300	127
1012	875	1046	159	634	4353	170	184	196	309	115
981	802	982	140	628	4190	158	178	172	278	100
901	682	947	109	598	4083	141	158	165	287	83
806	606	838	83	551	3972	121	138	153	267	66
812	580	767	63	542	3890	114	131	148	246	65
820	532	698	57	510	3602	106	117	129	234	57
1204	613	651	48	421	3606	96	107	122	223	68
1008	467	537	36	369	2798	69	98	83	157	53
766	350	438	28	301	2165	51	84	65	113	40
562	287	364	25	258	1797	41	83	49	108	31
432	287	387	27	244	1615	45	75	46	109	29
490	448	480	34	295	1598	81	106	84	331	46
678	603	743	88	384	2420	150	136	171	435	77
733	661	941	95	469	3281	153	149	189	203	99
682	681	839	112	495	3437	162	144	188	180	113
666	719	839	125	516	3779	168	154	208	311	129
700	821	798	131	547	4006	171	159	230	296	136
797	893	828	129	562	4274	177	166	245	300	139
778	795	913	125	525	4293	165	161	226	215	143
742	791	803	148	483	4013	140	157	214	199	130

St125	St109	St303	St311	St313	St315	St401	St801	St803	St805	St1201
823	883	857	128	540	4931	168	165	248	240	135
973	1003	1021	129	615	4990	182	169	268	186	143
1147	1189	1106	183	632	4616	176	176	285	222	134
1429	1462	1187	193	680	5596	207	172	277	270	146
1615	1345	1206	165	574	4888	184	168	246	305	137
1657	1052	1089	144	592	4754	165	157	218	268	125
1610	909	1068	113	548	4635	146	146	188	311	113
1396	818	953	90	507	4545	125	130	175	293	94
1273	763	843	68	504	4399	120	125	167	262	85
1250	730	763	61	471	4112	111	116	148	264	82
1659	563	689	66	491	3430	104	96	132	359	90
1371	399	567	52	411	2709	73	87	79	307	72
1040	263	462	44	325	2157	55	72	62	262	55
749	206	379	38	259	1805	42	65	38	225	43
557	227	434	33	237	1595	46	67	29	259	42
648	434	551	35	270	1493	78	92	95	315	74
890	733	851	79	427	2401	147	130	200	418	118
984	724	1054	90	518	3351	158	148	214	400	136
900	685	955	101	533	3379	171	142	218	374	150
857	753	962	111	536	3681	177	148	242	425	176
849	914	902	117	560	3736	177	153	271	533	190
921	1032	909	111	597	3841	180	164	292	438	194
871	912	1022	108	581	3740	167	154	265	607	199
773	920	879	138	497	3659	136	154	227	378	170
738	1038	914	124	590	4350	168	167	281	588	164
958	1100	1079	126	722	4427	186	170	314	464	167
1243	1227	1248	183	746	4336	177	204	337	628	152
1854	1369	1276	213	859	5157	228	175	318	433	165
2218	1256	1366	154	765	4334	198	167	296	496	159
2333	1047	1196	134	717	3310	172	155	264	409	150
2319	934	1189	116	702	4257	151	142	211	472	143
1986	839	1068	103	630	4249	129	129	197	472	122
1734	784	919	82	596	4092	126	120	186	434	105
1680	719	828	78	564	3827	116	108	167	795	107

VITAE

Mr. Syed Masiur Rahman earned his Bachelor of Science in Civil Engineering from Bangladesh University of Engineering & Technology (BUET) in June, 2000. He is a citizen of Bangladesh. He joined the Department of City & Regional Planning at King Fahd University of Petroleum & Minerals (KFUPM) in September, 2002 as a Research Assistant. He obtained Master of City & Regional Planning in June, 2004.

Mr. Syed is now serving the Department of Civil Engineering, KFUPM as a Lecturer-B and pursing his Ph.D. He is actively involved in many research projects of a few departments including department of Civil Engineering, Information and Computer Science, and Electrical Engineering, and Research Institute at KFUPM. Mr. Syed already published twelve (12) scientific articles in international referred journals. One of his articles is cited in a report of World Bank and IAIA. He is a co-author of two books, one of which is honored as a funded project by KFUPM.

Mr. Syed was awarded with a Certificate of Distinction-KFUPM (2005-2006), for being a member of the research team concerning the project titled "Environmental Impact Assessment – contract Area A in Northern Part of Rub Al-Khali, Stage 1-2D and 3D Seismic Operation (Project No. CEW 2290)."

Mr. Syed is a citizen of Bangladesh. He can be contacted through the following address.

Syed Masiur Rahman

Permanent Address

Village: Kanchana, P. O. Kanchana,
P. S. Satkania, Chittagong, Bangladesh

Present Address

946, West Sholoshahor, P. O. Amin Jute Mills
P.S. Panchlaish, Chittagong, Bangladesh.
Phone: +88-031-683070
E-mail: smrahman.bd@gmail.com